

Oxford
LINGUISTICS

The Noun Phrase

J. Rijkhoff

Oxford Studies in Typology and Linguistic Theory

The Noun Phrase

Jan Rijkhoff

Title Pages

(p.i) THE NOUN PHRASE

(p.ii) OXFORD STUDIES IN TYPOLOGY AND LINGUISTIC THEORY

(p.iii) THE NOUN PHRASE

SERIES EDITORS: Ronnie Cann, *University of Edinburgh*, William Croft, *University of Manchester* Anna Siewierska, *University of Lancaster*.

This series offers a forum for original and accessible books on language typology and linguistic universals. Works published will be theoretically innovative and informed and will seek to link theory and empirical research in ways that are mutually productive. Each volume will also provide the reader with a wide range of cross-linguistic data. The series is open to typological work in semantics, syntax, phonology, and phonetics or at the interfaces between these fields.

Published

Indefinite Pronouns

by Martin Haspelmath

Intransitive Predication

by Leon Stassen

Classifiers: A Typology of Noun Categorization Devices

by Alexandra A. Aikhenvald

Anaphora

by Yan Huang

The Noun Phrase

by Jan Rijkhoff

In preparation

Double Object Constructions

by Maria Polinsky

The Paradigmatic Structure of Person Marking

by Michael Cysouw

Subordination

by Sonia Cristaforo



(p.iv) *This book has been printed digitally and produced in a standard specification
in order to ensure its continuing availability*



Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research,
scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai
Nairobi
São Paulo Shanghai Taipei Tokyo Toronto

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© Jan Rijkhoff 202

The moral rights of the author have been asserted

Database right Oxford University Press (maker)

Reprinted 2004

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any
means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the
appropriate
regraphics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights
Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover

And you must impose this same condition on any acquirer

ISBN 0-19-823782-0

Contents

Title Pages

List of Figures

List of Tables

List of Abbreviations

Acknowledgements

1 Preliminaries

2 Nominal Subcategories: Seinsarten

3 Nouns: Real and Apparent Nominal Subclasses

4 Qualifying Modifiers in the Noun Phrase

5 Quantifying Modifiers in the Noun Phrase

6 Localizing Modifiers in the Noun Phrase

7 The Underlying Structure of Noun Phrases

8 Ordering Principles, Domain Integrity, and Discontinuity

9 Greenbergian Word Order Correlations and the Principle of Head Proximity

10 The Principle of Scope

11 Epilogue

References

Index of Subjects

Index of Languages

Index of Authors

(p.viii) List of Figures

- 1.1. Parts-of-speech systems 16
- 2.1. Cross-linguistic classification of major nominal subcategories 54
- 3.1. Stages in the acquisition of class/gender or noun markers 93
 - 4.1. Parts-of-speech systems 122
 - 7.1. Tense 215
- 7.2. Symmetry in the underlying structure of the clause and the NP 224
 - 7.3. The hierarchical structure of the NP 226
 - 9.1. Parts-of-speech systems 288
- 9.2. Noun modifiers in the fifty-two-language sample and the VO/OV distinction 303
 - 11.1. Cross-linguistic classification of major nominal subcategories (or *Seinsarten*) 338
 - 11.2. The hierarchical structure of the NP 343
 - 11.3. The hierarchical structure of the clause (predication) 344

(p.ix) List of Tables

1.1.	Distribution of sample languages over (sub)phyla (Ruhlen's classification)	6
1.2.	Sample languages and their locations	9
1.3.	Parts-of-speech systems of languages in the sample	18
2.1.	Nominal subcategories in sample languages: singular object nouns, set nouns, and sort nouns	30
2.2.	-Shape and +Shape nominal subcategories	51
2.3.	Verbal categories	58
3.1.	Properties associated with prototypical exemplars of classifier and noun class systems	74
3.2.	Gender/number in Krongo	86
3.3.	Some declension classes in Latin	96
4.1.	Basic aspectual categories of the verb	104
4.2.	Languages with set nouns	107
4.3.	Adjective-noun order in languages of types 1–3: V/N/A, V-N/A, V-N-A	129
4.4.	Adjective-noun order in languages of intermediate type 3/4: V-N(-A)	133
4.5.	Parts-of-speech systems of languages in the sample	142
5.1.	Nominal subcategories and reduplication	148
5.2.	Noun-numeral constructions	158
6.1.	The expression of attributive demonstratives	179
6.2.	The expression of (in)definiteness	186
6.3.	Localizing satellites in the NP	195
7.1.	Basic aspectual categories of the verb	216
7.2.	Discourse operators Ω and Π	235
8.1.	Some typological data of sample languages	240
9.1.	Classification of languages	261
9.2.	Exceptions to Hawkins's universals	277
9.3.	Dryer's correlation pairs	279
9.4.	Dryer's correlation pairs	298
9.5.	Word order data of sample languages (not included: Cayuga, Etruscan, Meroitic)	300
9.6.	The relation between the position of adjective, demonstrative, and numeral	304
9.7.	Basic word order and the position of the possessor pronoun relative to the noun	306
	(p.x)	
9.8.	Possessor affixes and possessor NPs (G) in twenty languages	306
9.9.	Free possessor pronouns (g) and possessor NPs (G) in twenty-nine languages	307
10.1.	Constituent order in the simple NP	328

(p.xi) List of Abbreviations

Examples from different languages are given in the transcription of the original source.

- :
separates lexical and grammatical categories
- .
separates grammatical categories
- =
marks boundary between morpheme and clitic
- –
separates morphemes
- 1
first person
- 2
second person
- 3
third person
- A
adjective
- ABL
ablative
- ABS
absolutive
- ACC
accusative
- ACT
actual
- ActFoc
actor focus
- ADV
adverb
- ANAPH
anaphoric pronoun
- anim
animate
- ANPA
anti-passive
- AOR
aorist
- AP
attributive particle
- ART
article
- ASP
aspect
- AUX
auxiliary verb
- AUX1
intransitive auxiliary

- AX
affix
- BEN
beneficiary
- C
common gender
- CL
class
- CLF
classifier
- CN
connector
- COLL
collective
- COMP
complementizer
- CONJ
conjunction
- CONT
continuative
- COP
copula
- COR
correlative marker
- CV
converb
- CX
circumfix
- DAT
dative
- DECL
declarative
- DEF
definite
- DEM
demonstrative
- DIM
diminutive
- DistPC
distant past continuative
- DO
direct object
- DS
different subject
- DU
dual
- DUR
durative
- ERG
ergative

- ES
ergative suffix
- EXH
exhortative
- F
feminine
- FIN
finite
- FOC
focus
- FUT
future
- GEN
genitive
- GENR
general tense-aspect-mood marker
- H
high tone
- HAB
habitual
- HSY
hearsay
- HUM
human
- IMP
imperative
- IMPF
imperfective
- IN
inclusive
- INDEF
indefinite
(p.xii)
- IND/NML
indicative/nominalization
- INGR
ingressive
- INST
instrumental
- INTR
intransitive
- IO
indirect object
- IRR
irrealis
- L
low tone
- LD
locative-directional
- LIN=GEN

linker clitic functioning as a genitive

- LIN-SEP
linker clitic plus separating element (indicating scope)
- LOC
locative
- LVP
locative verbal prefix
- M
masculine
- MLOC
modal locative
- N
neuter
- NLZR
nominalizer
- NM
noun phrase marker
- NOM
nominative
- NomP
nominalizing particle
- NonF
nonfinite
- NonH
nonhuman
- NSP
non-specific
- NUM
numeral
- O
object
- OBL
oblique
- obl
obligatory
- opt
optional
- PART
participle
- PassPart
passive participle
- PAST
past
- PastPart
past participle
- PERF
perfective
- PL
plural
- PNCT

- punctual
- POS
 - possessive
- PP
 - past punctual
- PRES
 - present
- PRET
 - preterite
- PROP
 - proprietary
- PROX
 - proximate
- PTM
 - phrase terminal marker
- PX
 - prefix
- QM
 - question marker
- RC
 - relativizing case suffix
- RED
 - reduplication
- REFL
 - reflexive
- REL
 - relative clause marker
- RelP
 - relative particle
- RemP
 - remote past
- RSM
 - resumptive marker
- S
 - subject
- SC
 - subject concord
- SER
 - serial
- SG
 - singular
- SING
 - singulative
- SP
 - specific
- S.Part
 - subject participle
- SR
 - same referent
- SS

same subject

- STAT
stative
- SubSx
subordinative suffix
- SubV
subordinate verb form
- SUP
supposition
- SUPP
support verb
- SX
suffix
- TN
transitional nasal
- TNS
tense
- TOP
topic
- TR
transitive
- V
verb
- VAL
validator

(p.xiii) Acknowledgements

A typology book cannot be written without help from others. I am indebted to the following, who supplied me with information: Dik Bakker, Peter Bakker, Machtelt Bolkestein, Lyle Campbell, G. Tucker Childs, Bernard Comrie, Milly Crevels, Michael Fortescue, Casper de Groot, J. Houwink ten Cate, Peter Kahrel, Ankie Klootwijk, Maria Koptjevskaja-Tamm, Silvia Kouwenberg, Manfred Krifka, Chryssoula Lascaratou, Silvia Luraghi, William McGregor, J. Lachlan Mackenzie, Johanna Mattissen, Edith Moravcsik, Pieter Muysken, Vladimir P. Nedjalkov, Johanna Nichols, Michael Noonan, Hella Olbertz, Harm Pinkster, Frans Plank, Mechthild Reh, Albert Rijksbaron, Tapani Salminen, Hans-Jürgen Sasse, Gerjan van Schaik, Christoph Schroeder, Johanna Seibt, Anna Siewierska, Rienk Smeets, Arie Spruit, Stanley Starosta, Leon Stassen, Maria Vilkuna, Hein van der Voort, Lourens de Vries, Gerry Wanders, Heinrich Werner, and Elizabeth Zeitoun. I am especially grateful to Edith Moravcsik and Anna Siewierska for their invaluable comments.

The work of three people has shaped the contents of this book in particular: the late Simon Dik (who also supervised my Ph.D.), the late Joseph Greenberg, and Kees Hengeveld. Dik's approach to languages and grammatical theory made me decide to become a linguist; Greenberg's article on universals of language (Greenberg 1966a) triggered a lasting fascination with linguistic typology; and Hengeveld's ideas concerning the layered representation of clausal structures in Dik's *Theory of Functional Grammar* have provided me with a highly inspiring framework for the study of noun phrases.

Finally, I am most grateful to my parents and my family. My parents have always provided me with unlimited support; sadly, my mother did not live long enough to see the publication of this book. My wife and daughter have contributed to this book in various ways, enriching it with dimensions that perhaps only I can see.

J.R.

Beder

September 2001 (p.xiv)

1 Preliminaries

1.1. Introduction

This chapter discusses some general properties of noun phrases and sets the framework for subsequent chapters. Section 1.2 states the aims and gives an outline of the organization of this book. The theoretical background is motivated in section 1.3, which also introduces the three notions that are used to characterize most of the noun modifiers: Quality, Quantity, and Location. Since this study is largely based on data from a representative sample of fifty-two languages, the next section is concerned with the sampling method and section 1.5 discusses such central notions as *noun*, *noun phrase*, and *referent* and gives an overview of the cross-linguistic variation with respect to the internal structure of the noun phrase (henceforth NP).

1.2. Aims

This book has three aims. First of all it seeks to provide a cross-linguistic account of the constituents of the noun phrase. In spite of the fact that languages do not necessarily make use of the same kind of noun (for the same kind of thing) or the same set of noun modifiers (e.g. many languages lack articles and/or adjectives), this is not an issue that has received considerable attention in linguistic typology. Yet we can only establish the cross-linguistic characteristics of NPs after we have determined the actual existence of the constituents whose properties we are investigating. Thus, since each individual language only uses a subset of all the possible categories, forms, and constructions, it is necessary first to investigate the noun and to establish the range of grammatical and lexical noun modifiers in the world's languages. This is the main topic of Chapters 2–6, in which I will present an overview of the various NP-internal categories and their semantic and morpho-syntactic properties. The information contained in these chapters is especially relevant in connection with the word order principles that are discussed in Chapters 8–10, which are devoted to the internal syntax of NPs. For instance, since the word order principles that I propose apply to *free* constituents of an *integral* (or *hierarchical*) NP, we need to establish for each language in the sample whether noun modifiers are indeed free constituents, and whether they are in a direct construction with the head noun (and not in apposition, or constituting a distinct phrase at the level of the clause).

(p.2) The second aim of this book is to present a typologically adequate model of the (integral) NP in the general framework of Simon Dik's *Functional Grammar* (Dik 1997), i.e. a formal representation of the underlying NP structure that is designed to accommodate NPs from any type of natural language. I will argue that the NP can be analyzed as a layered structure in which the various kinds of NP operators (grammatical noun modifiers such as determiners and quantifiers) and NP satellites (lexical noun modifiers such as adjectives, possessor phrases, relative clauses) can be associated with one of four hierarchically nested layers in the underlying structure. I will also contend that, at some level of abstraction, the underlying structure of the clause (or, rather, predication) is similar to that of the NP.

The third aim of this study is to account for word order patterns in NPs as they are attested in the languages of the world. In Chapters 8–10 a small number of universal ordering principles are proposed and tested against data taken from the fifty-two-language sample. These chapters are primarily concerned with the internal syntax of *simple* NPs, i.e. with the relative order of the determiner, numeral, adjective, and noun. The main reason for this restriction is that there is hardly any information in the literature about the position of embedded modifiers such as the possessor phrase and the relative clause vis-à-vis other constituents of the simple NP. Everyone familiar with cross-linguistic research knows that many grammars are not very informative about the relative order of constituents in simple NPs, let alone complex NPs. As a matter of fact, quite a number of grammars do not discuss the internal syntax of NPs at all.

Even within the simple NP not every possible word category will be taken into account, but only the demonstrative pronoun, the cardinal numeral (plus numeral classifier, if any), the adjective, and the head noun. This is due to the fact that, if in a grammar attention is paid to the internal syntax and semantics of NPs, it is usually only these categories that are considered. Also, in almost every language a demonstrative, a cardinal numeral, an adjective, and a head noun can appear in the same NP (at least in those languages that have these four categories), which makes it easier to set up a cross-linguistic investigation.

In short, this study has three aims. First it wants to give a cross-linguistic overview of the semantic and morpho-syntactic properties of the constituents of the NP (Chapters 2–6). Secondly, I want to propose a new, typologically adequate model of the underlying structure of the NP (Chapter 7), and in the remaining chapters (8–10) I will demonstrate that NP-internal ordering patterns can be accounted for in terms of a few general ordering principles.

1.3. The theoretical framework

Most facts and analyses in this book are presented in a more or less theory independent fashion. The reason why I have chosen to avoid as much as possible ([p.3](#)) theory specific formulas, symbols, and representations is simply because I wanted this book in the first place to be a book about noun phrases that would be of interest for linguists of any theoretical persuasion.

On the other hand I have not tried to hide the fact that this study was conducted within the general framework of Dik's theory *Functional Grammar* (Dik 1997). Every linguistic researcher uses some kind of theory, whether this is just a set of hypotheses concerning the phenomenon under investigation or one of the more sophisticated general theories of grammar such as Dik's *Functional Grammar*, Van Valin's *Role and Reference Grammar*, a version of Chomsky's theory, or some other theoretical framework. Even when one is interested in such a seemingly simple problem as the cross-linguistic distribution of the indefinite article, one must have a set of more or less coherent, well-defined ideas about indefinite articles, for in order to be able to study indefinite articles, one must at least be able to recognize them (e.g. by function, form, meaning, position, or a combination of these things). In other words, one must have a theory about the way they can be identified.

There are several reasons why I chose to conduct my investigation in the larger context of Dik's theory of *Functional Grammar* (henceforth: FG). Perhaps the most important

reason for adopting the theory of FG is that I share Dik's basic assumption that languages should be studied in the light of their communicative function. For me not to do so would be comparable to trying to understand the form and structure of, say, cars without somehow taking into account the fact that they are designed to provide a means of transport.¹ This is not to say that everything in linguistics can or must be explained in functional terms. The fact that natural languages are a species specific phenomenon implies that genetic factors are also involved. But as long as the exact nature and impact of the genetic component remains far from clear, it seems sensible only to resort to genetic explanations for linguistic facts when it is impossible to find an explanation on the basis of less controversial evidence from other areas such as semantics, pragmatics, or cognitive psychology. Moreover, genetic factors might appear to be explainable in functional terms as well (Dik 1986, 1997; cf. also Croft 1991: 1; Nuyts 1992: 77).

Another important reason for adopting the FG framework is related to the fact that in this theory the underlying structure is a *semantic* rather than a syntactic representation, in which elements are not linearly ordered (which is not to say that there is no structure in the underlying FG representation). I do not believe (**p.4**) there is a reason for postulating some underlying order of constituents in a general theory of grammar and to my knowledge the psychological reality of a linearly ordered underlying structure has never been proven (cf. Hawkins 1983: 125; Dik 1997: 393). Last but not least, FG is one of very few general theories of grammar that takes languages seriously in that it systematically takes into account typological facts from a wide variety of languages.

For the reasons mentioned above it will hardly come as a surprise that the underlying structure of the NP that I put forward in Chapter 7 can most easily be incorporated in the Theory of Functional Grammar as originally developed by Simon Dik (see Dik 1997: 161 f., 402). But Van Valin has demonstrated that (an earlier version of) the layered model of the NP put forward in this book can also be fitted into his Role and Reference Grammar (Van Valin and LaPolla 1997: 52–9), and Craig (1992) used the same model to handle various types of classifiers (see also Grinevald 2000: 76–7).

In Chapters 4–6 noun modifiers are characterized in terms of three notions: Quality, Quantity, and Location. I have taken these distinctions from Aristotle who wrote in *Physics* (5.2): '*movement or change ... pertains exclusively to quality, quantity, and locality, each of which embraces contrasts*' (see also Rijksbaron 1989). Although Aristotle was concerned with physics rather than linguistics I hope to demonstrate that the function of both grammatical and lexical modifiers of the noun (as well as verbs) is aptly captured by these names.

Let me finally point out that this study does not offer detailed analyses of NPs as they occur in the fifty-two languages of the sample.² Instead data from these languages are used to construct a typologically adequate model of the underlying structure of the NP and to test hypotheses concerning word order patterns. In sum, besides many primary data this study offers:

- the introduction of two new grammatical categories, namely nominal aspect (Chapter 4) and *Seinsart* (Chapter 2);
- new universals (e.g. Chapter 4 on the occurrence of adjectives as a distinct word class);

(p.5)

- a new, layered model of the underlying structure of the NP, as well as possible parallels with the underlying structure of the clause (or, rather, predication—Chapter 7);
- an account of word order patterns in the simple NP in terms of a limited set of general ordering principles (Chapters 8–10).³

1.4. The sample

This study is based on data from a representative sample of fifty-two languages that were selected on the basis of a sampling method developed by Rijkhoff et al. (1993). This method is not so much designed to suit statistical or probabilistic purposes, but rather to reveal as much as possible about the range of linguistic variation in the languages of the world. The underlying assumption is that the best way to achieve this is via genetic diversity. To account for linguistic diversity *across* maximal genetic groupings (or phyla) the sampling procedure requires that every known phylum is represented by at least one language. To account for linguistic diversity *within* phyla remaining languages must be proportionally distributed over the phyla on the basis of a factor called Diversity Value (DV). This value indicates the linguistic diversity in a group of genetically related languages: the greater the linguistic diversity of a language family, the higher the DV and the more languages of that family will be included in the sample. DVs are calculated on the basis of the internal make-up of the phylum, the basic assumption being that the linguistic diversity in a group of genetically related languages correlates with historical relations between languages. The more branches a phylum has close to the top node of the tree, the more diverse it is taken to be. DVs are also used to distribute languages over subphyla (and sub-subphyla etc.); the larger the sample the more internal subgroupings of each phylum are represented by at least one language. Recursive application of the procedure in a top-down fashion ensures that in the end the genetic distance between individual sample languages is always maximal (for a detailed description of the whole procedure and a discussion of other sampling methods, I refer the reader to Rijkhoff et al. (1993) and Rijkhoff and Bakker (1998)).

Using Ruhlen's classification of the world's languages (Ruhlen 1987, 1991), this resulted in the selection shown in Table 1.1 (the numbers in round brackets indicate the DV, the number of subphyla, and the number of languages; for example, the DV of Afro-Asiatic is 55.53, it divides into six subphyla, and the family contains 258 languages). Since there are more sample languages than phyla (at least in Ruhlen's classification; he recognizes twenty-nine phyla, (p.6)

Table 1.1. Distribution of sample languages over (sub)phyla (Ruhlen's classification)

Family	No. of lgs. in sample	Language
Afro-Asiatic (55.53/6/258)	2	
Chadic (19.18/4/123)		Gude
Cushitic (9.04/2/36)		Oromo
Altaic (15.07/2/62)	1	Turkish
Amerind (178.44/6/854)	7	
Central Amerind (19.05/3/60)		Pipil

Ge-Pano-Carib (29.25/2/193)		Hixkaryana
Northern Amerind (45.48/3/232)	(2)	
Almosan-Keresiouan (15.98/2/97)		Cayuga
Penutian (21.46/8/92)		Koasati
Equatorial-Tucanoan (44.96/2/268)		Guaraní
Chibchan-Paezan (16.91/2/71)		Ika
Andean (9.50/6/30)		Imbabura Quechua
Australian (67.58/30/262)	3	
Gunwiny gun (6.50/6/11)		Ngalakan
Pama-Nyungan (39.25/30/195)		Kayardild
Nunggubuyu (0.00/0/1)		Nunggubuyu
Austric (137.41/3/1186)	5	
Austro-Tai (106.03/2/1027)	(3)	
Austronesian (118.17/4/970)	(2)	
Tsouic (2.45/2/4)		Tsou
Malayo-Polynesian (131.05/2/950)		Samoan
Daic (4.67/2/57)		Nung
Austroasiatic (28.08/2/155)		Vietnamese
Miao-Yao (2.00/2/4)		Miao (or: Hmong)
Basque (language isolate)	1	Basque
Burushaski (language isolate)	1	Burushaski
Caucasian (5.98/2/34)	1	Abkhaz
Chukchi-Kamchatkan (2.47/2/5)	1	Chukchi
Elamo-Dravidian (7.43/2/29)	1	Tamil
Eskimo-Aleut (3.34/2/9)	1	West Greenlandic
Etruscan (language isolate)	1	[Etruscan]
Gilyak (language isolate)	1	Gilyak
Hurrian (language isolate)	1	Hurrian
Indo-Hittite (39.71/2/180)	2	
Indo-European (36.94/9/175)		Dutch
Anatolian (4.00/4/5)		Hittite
Indo-Pacific (123.39/13/748)	5	
Trans-New Guinea (66.89/21/508)		Wambon
Sepik-Ramu (15.25/5/98)		Alamblak
West Papuan (5.33/4/24)		Galela
Torricelli (9.50/7/48)		Bukiyip
East Papuan (6.89/3/27)		Nasioi

Kartvelian (3.00/3/4)	1	Georgian
Ket (language isolate)	1	Ket
Khoisan (6.97/3/33)	1	Nama Hottentot
Korean-Japanese-Ainu (3.00/3/4)	1	Korean
Meroitic (language isolate)	1	[Meroitic]
Nahali (language isolate)	1	Nahali
Na-Dene (9.44/2/41)	1	Sarcee
Niger-Kordofanian (90.38/2/1068)	4	
Niger-Congo (90.07/2/1036)	(3)	
Niger-Congo Proper (89.68/2/1007)	(2)	
Central Niger-Congo (91.16/2/961)		Babungo
West Atlantic (10.05/3/46)		Kisi
Mande (9.30/3/29)		Bambara
Kordofanian (9.51/2/32)		Krongo
Nilo-Saharan (42.18/9/138)	2	
East Sudanic (20.82/4/80)		Lango
Central Sudanic (7.20/2/36)		Ngiti
Pidgins and Creoles (13.47/13/38)	1	Berbice Dutch Creole
Sino-Tibetan (38.52/2/268)	2	
Sinitic (4.10/2/13)		Mandarin Chinese
Tibeto-Karen (28.71/2/255)		Burmese
Sumerian (language isolate)	1	Sumerian
Uralic-Yukaghir (4.92/2/27)	1	Hungarian

(p.7) including nine singleton phyla, i.e. language isolates, and a group of Pidgin and Creole languages), this fifty-two-language sample contains thirty-two languages with a shared genetic membership.⁴ These languages were distributed proportionally over the phyla on the basis of their DVs in a top-down fashion according to the procedure briefly explained above (first phyla, then subphyla, then sub-subphyla, etc.). In other words, (sub)phyla with a relatively high DV are assigned(p.8) more languages than (sub)phyla with a relatively low DV. When there are more subphyla than languages to be distributed, the method requires that each language belongs to a different subphylum. For instance, the Afro-Asiatic phylum divides into six subphyla (Ancient Egyptian, Berber, Chadic, Omotic, Cushitic, and Semitic), but the two Afro-Asiatic languages in the sample may not both belong to the same subphylum.

In those cases where the languages outnumber the subphyla, the sampling procedure requires that the procedure is repeated and that languages are distributed over the subphyla on the basis of DV so that subphyla with high DVs are better represented than subphyla with relatively low DVs. This occurs, for instance, in the Amerind phylum where seven languages must be distributed over six subphyla: Northern Amerind, Central Amerind, Chibchan-Paezan, Andean, Equatorial-Tucanoan, and Ge-Pano-Carib. First each subphylum gets one language, then the remaining seventh language is

assigned to the subphylum that has the highest DV, namely Northern Amerind. Since this subphylum divides into three groups (sub-subphyla; Almosan-Keresiouan, Penutian, Hokan), the sampling procedure requires that the two Northern Amerind languages in the sample must belong to different subfamilies. The procedure outlined above was repeated until all fifty-two sample languages were distributed on the basis of DV values. The actual sample, however, contains forty-nine languages because hardly anything is known about Nahali and the extinct language isolates Etruscan and Meroitic.⁵ So as not to distort the proportions within the sample, they have not been replaced by other languages.

Table 1.2 gives the approximate locations where the sample languages are spoken; the information is largely based on information contained in Grimes (1988). If a nation or region is in italic print, this means that the grammar that I used for this study describes the variant spoken in that particular nation or region. Note finally that I use the usual current name *Hmong* rather than *Miao* (the dialect that I used for this book is called *Hmong Njua*, lit. *Green or Blue Hmong*), and *Nivkh* rather than *Gilyak*.

1.5. Nouns, noun phrases, and referents

This section is concerned with such general questions as: ‘Do all languages have nouns and (integral) NPs?’ and ‘What do NPs refer to?’⁶ I will show that there are languages in which a major distinct class of nouns appears to be lacking and that many languages have nouns that cannot be distinguished from other major word classes (verbs, adjectives). I will also demonstrate that even if a language (p.9)

Table 1.2. Sample languages and their locations

Languages in the sample	Approximate location
1. Abkhaz	Georgia (Caucasus)
2. Alamblak	Papua New Guinea (East Sepik Province)
3. Babungo	Cameroon (North West Province, Ndop plain)
4. Bambara	Mali
5. Basque	Spain, France (both sides of Western Pyrenees)
6. Berbice Dutch Creole	Guyana
7. Bukiyp (Mountain Arapesh)	Papua New Guinea (East Sepik Province, Prince Alexander Range)
8. Burmese	Burma
9. Burushaski	Pakistan
10. Cayuga	Canada (Ontario), USA (New York, Oklahoma)
11. Chukchi	Russia (NE Siberia, Chukchi peninsula)
12. Dutch	The Netherlands, Belgium (Flanders)
13. [Etruscan]	Present-day Italy
14. Galela	Indonesia (Halmahera)
15. Georgian	Georgia (Caucasus)
16. Gilyak = Nivkh	Russia (Sakhalin Island)
17. Greenlandic, West	Greenland

18. Guarani	Paraguay, Argentina, Bolivia, Brazil
19. Gude	Nigeria, Cameroon (Mokolo and Guider Provinces)
20. Hittite	Present-day Turkey (Anatolia)
21. Hixkaryana	Brazil (Northern, Nhamundá, and Mapuera river area)
22. Hungarian	Hungary
23. Hurrian	Present-day southern Turkey
24. Ika	Columbia (Sierra Nevada de Santa Marta)
25. Kayardild	Australia (Bentinck Island)
26. Ket	Russia (Siberia, Yenisey river area)
27. Kisi	Guinea (Guéckédou and Kissidougou districts), Sierra Leone, Liberia
28. Koasati	USA (south-west Louisiana)
29. Korean	Korea
30. Krongo	Sudan (southern part of Nuba Mountains)
31. Lango	Uganda (Lango Province)
32. Mandarin Chinese	China
33. Meroitic	Present-day Sudan
34. Miao = Hmong	South-East Asia
35. Nahali	India (Nimar District)
36. Nama Hottentot	Namibia, South Africa
37. Nasioi	Papua New Guinea (North Solomons Province, Bougainville)
38. Ngalakan	Australia (Arnhem Land)
39. Ngiti	Zaire (Haut-Zaire Région)
40. Nung	Vietnam
41. Nunggubuyu	Australia (Northern Territory)
42. Oromo	Kenya, Ethiopia
43. Pipil	El Salvador
44. Quechua, Imbabura	Ecuador, Argentina, Bolivia, Chile, Peru
45. Samoan	Samoa
46. Sarcee	Canada (Alberta)
47. Sumerian	Present-day Iraq
48. Tamil	India (Tamil Nadu, North Arcott district), Sri Lanka
49. Turkish	Turkey
50. Tsou	Taiwan (Alischan region)
51. Vietnamese	Vietnam
52. Wambon	Indonesia (Southern Irian Jaya)

(p.10) does employ nouns, this does not necessarily mean that the language in question also has NPs. Section 1.5.5, finally, is devoted to the status of referents of NPs.

I will argue that referents of NPs are mental constructs whose properties need not coincide with the properties of their ontological correlates in the external world (if they exist).

1.5.1. Word classes

Despite many attempts in the recent and not so recent past to provide definitions of the major lexical word classes, it is safe to say that there is still no general consensus among typologists on what constitutes a verb, a noun, or an adjective.⁷ This is mostly due to the fact that it has proven to be rather difficult to define word classes in a language independent fashion. For example, to say that a noun is a word that is inflected for number is quite irrelevant for all those languages across the globe in which number marking is absent (see Chapter 5). Recently, however, Hengeveld (1992a) and Croft (1990b, 1991, 2000) have proposed definitions that avoid these problems.

In Hengeveld's approach crucial reference is made to the function(s) that a lexical item can fulfill without having to resort to special grammatical measures. Thus one can distinguish between lexical items that (without any special measures) can be used as the head of a referring expression (NP or term phrase), (p.11) i.e. nouns, and lexical items that can be used (without having to resort to any special measures) to modify the head of a term, i.e. adjectives (more on this in this section 1.5.2 below).

Croft on the other hand argues that verb, noun, and adjective are not categories of particular languages but rather language universals in the sense that they constitute what he calls 'typological prototypes'. Instead he focuses on the constructions that are used for the three universally attested communicative functions of predication, reference, and modification. In his view (Croft 2000: 84–5, 87),

Categories in a particular language are defined by the constructions of the language. Moreover, the constructions are the primitive elements of syntactic representation; categories are derived from. ... constituents Constructions define grammatical categories.

The range of constructions in the universal-typological theory of parts of speech covers constructions for predication, reference and modification. Most important, it explicitly recognises that predication, reference and modification are pragmatic (communicative) functions or, as Searle described them, propositional acts.

Croft then argues that one should focus on the unmarked combination of pragmatic function and lexical class and that in the case of parts-of-speech the unmarked combinations are (Croft 2000: 88):

lexical class	pragmatic function
noun	reference to an object
adjective	modification by a property
verb	predication of an action

Any other combination of pragmatic function and lexical class is marked. Thus, an object word such as 'vehicle' is unmarked for reference but marked for predication ('be

a vehicle') and modification ('vehicular, vehicle's'; cf. Palmer 1992). Croft's approach has much in common with the ideas proposed earlier by Hopper and Thompson (1984), who investigated properties of word classes from a discourse perspective. They stated that the basic categories N and V are best viewed as 'universal lexicalizations of the prototypical discourse functions of "discourse-manipulable participant" and "reported event", respectively' and concluded that 'categoriality itself is another fundamental property of grammars which may be directly derived from discourse function' (Hopper and Thompson 1984: 703).

In this book I will follow Hengeveld's approach to parts-of-speech systems (section 1.5.3), not only because he stays closer to the linguistic data (as will be shown below, the parts-of-speech system he recognizes closely reflects statements and data provided in the grammars), but also because he offers a *typology* (rather than just a classification) of parts-of-speech systems in that it appears to be possible to predict certain semantic or morpho-syntactic features of a language (p.12) once one knows what kind of parts-of-speech system that language employs (cf. Rijkhoff 2000; Hengeveld and Valstar forthcoming; Hengeveld et al. forthcoming).

1.5.2. Nouns

A cross-linguistic investigation of the issues that are raised in connection with the establishment of nouns as a word class in its own right reveals that languages can be divided into three types:

1. languages without a major word class of nouns;
2. languages in which nouns cannot be distinguished from other word classes (verbs, adjectives);
3. languages with a distinct class of nouns.

There is some controversy over the question whether there really are languages without nouns, but experts seem to agree that in Cayuga, Tuscarora, Mohawk, and other Iroquoian languages nouns are at best a minor word class. Sasse (1993) has argued that Cayuga does not have nouns but only verbs, which can be divided into two subcategories. The first group consists of a small number of uninflected and monomorphemic forms which Sasse calls *Simplizia*. To this subcategory belong a number of animal names (originating mostly as ideophones), a few loanwords, and words of unknown origin that are used for objects: *twę:twę:t* 'duck', *ka ka:* 'crow', *kwihskwihš* 'pig', *kwęńihs* 'pennies', *khá:fih* 'coffee', *só:ta:* 'soldier'.

The second subcategory contains a large group of roots (*Wurzeln*) that occur with at least one pronominal prefix and an aspect suffix. Sasse (1993: 206) distinguishes between two kinds of roots: R1 and R2 roots. R1 roots normally only appear with one pronominal prefix (usually the third person singular non-human form) and a stative aspect suffix. They are largely used to refer to discrete physical objects, e.g.: *ka-nhóh-a'* 'it is a door' (/-*nhoh-*/ '[be a] door'), *ka-nyó: t-a'* 'it is a spoon' (/-*nyot-*/ '[be a] spoon'). R2 roots, on the other hand, can occur with all pronominal, tense, aspect, and mood affixes as well as with other kinds of affixes, e.g. *ha-hyátq-ha'* 'he writes it (down)' (/-*hyátq-*/ 'write'), *a-k-á:tkęh* 'I stood up' /-*á:tkęh-*/ 'stand up'), *o-yá:nr-e'* 'it is good' (/-*yanr-*/ 'be good').⁸

In short, in Sasse's view speakers of Cayuga commonly refer to an object by means of a phrase whose nucleus consists of a R1 root, which is basically a verbal predicate (Sasse 1993: 209). The only difference between R1 and R2 roots is that the former occur with a subset of the person and tense/aspect/mood affixes ([p.13](#)) that can appear on R2 roots. Sasse concludes that Cayuga does not have a lexical category that can be characterized as nouns.⁹

Below are some examples from Tuscarora, another Iroquoian language:

Tuscarora (Mithun Williams 1976: 29–30, 32)

(1)	yekhwaráhkhwá?		
	ye-khw-a-r-ahkw-ha?		
	HUM-food-JOINER-in-INS-SER		
	one_uses_it_for_having_food-in ('stomach')		
(2)	rò:rá:thv:		
	r-o-rathv-		
	M-OBJECTIVE-climb-PERF		
	he_climbs ('black snake')		
(3)	rakwá:tihs	wahratkáhtho?	katéskrahs
	ra-kwatihs	wa-hr-at-kahtho-?	ka-teskr-ahs
	M-young	AOR-M-look_at-PNCT	NonH-stink-SER
	he_is_young	he_looked_at_it	it_stinks
	('the boy looked at the goat')		

In her analysis of Tuscarora NPs, Mithun Williams (1976: 31) takes a less extreme position than Sasse regarding the verb/noun problem in Iroquoian languages by distinguishing between form and function (see also Mithun 2000).¹⁰ While admitting that 'all of these descriptive labels [in Tuscarora examples above—JR] are surface morphological verbs', she also claims that 'they clearly function in the same way as formal nouns syntactically. They occur with determiners, for example, and they can be conjoined with formal nouns to function as compound subjects or objects. The fact that many noun phrases are actually realized as surface verbs, while they function just as common nouns, provides additional support for the analysis of nouns as semantic propositions.'¹¹

Samoan also seems to have only one major word class, but it differs from Cayuga and Tuscarora in that its members are extremely flexible; they may be ([p.14](#)) used as a verb, a noun, or an adjective. Consider, for instance, the following citation from Mosel and Hovdhaugen (1992: 73, 74, 77).¹²

Many, perhaps the majority of, roots can be found in the function of verb phrase and NP nuclei and are, accordingly, classified as nouns and as verbs. This does not mean that a noun can be used as a verb or a verb as a noun or that we have two

homophonous words, one being a noun and the other being a verb. Rather, it means that in Samoan the categorization of full words is not given a priori in the lexicon. It is only their actual occurrence in a particular environment which gives them the status of a verb or a noun.... What is given in the lexicon, is not a particular word class assignment, but the potential to be used in certain syntactic environments as a noun or a verb. (p. 77)¹³

Although certain full words seem to be used more as verb or more as an NP nucleus for semantic reasons, there are no lexical or grammatical constraints on why a particular word cannot be used in the one or the other function. (p. 73)

Most Samoan equivalents of English adjectives, particularly the typical ones, are full words which function not only as attributes, but also as a noun or verb phrase nucleus, e.g. *lelei* '(be) good', *leaga* '(be) bad'; and conversely, all full words which function as noun and verb phrase nucleus can also be used as attributive modifiers, (p. 74)¹⁴

Here are some examples of roots with their verbal and nominal translations in English:

Samoan (Mosel and Hovdhaugen 1992: 73 f., 82 f.)

	noun phrase nucleus	verb phrase nucleus
(4) a.	teine 'girl'	'be a girl'
b.	tusi 'book, letter'	'write'
c.	salu 'broom'	'sweep'
d.	ma'i 'patient, sickness'	'be sick'
e.	la 'sun'	'be sunny'
f.	fana 'gun'	'shoot'
g.	lama 'torch'	'fish by torch light'

It is basically the presence of non-lexical elements that indicates what particular function such predicates fulfill. If a flexible predicate serves as the head of the (p.15) clause, it will typically combine with tense-aspect-mood particles; if it serves as the head of the term it will appear with an article or a preposition.

Tongan is another example of a language with extremely multifunctional predicates.¹⁵ This is shown in the following examples where the word *si'i* '(to be) small, smallness' is used as a verb in (5), as a noun in (6), and as an adjective in (7).¹⁶

Tongan (Tchekhoff 1981: 4)

(5) Na'e si'i 'ae akó
PAST small ABS school:DEF
'The school was small'
(6) 'i 'ene si'i

in POS.3SG childhood:DEF

'in his/her childhood'

- (7) Na'e ako 'ae tamasi'i si'i iate au
PAST study ABS child little LOC 1SG

'The little child studied at my house'

Whereas Samoan and Tongan have a single class of lexemes whose members combine the prototypical functions of verb, noun, and adjective (V/N/A), Quechua is said to have two major word classes: a distinct class of verbs and a large class of words which 'includes what in other languages would be distinguished as nouns and adjectives. These are regarded as a single class ... because there is insufficient evidence of a strictly morpho-syntactic nature for distinguishing them (as lexical categories)' (Weber 1989: 35). Examples (8)–(11) show that the Quechua counterparts of the English noun 'mayor' *alkalde* and the English adjective 'big' *hatun* can serve as a noun, as in (8) and (10), and as an adjective, as in (9) and (11). Compare:

Quechua (Schachter 1985: 17)

- (8) Rikaška: alkalde-ta

see:PAST.1SG mayor-ACC

'I saw the mayor'

- (9) chay alkalde runa

DEM mayor man

'that man who is mayor'

- (10) Rikaška: hatun-ta

see:PAST.1SG big-ACC

'I saw the big one'

(p.16)

- (11) chay hatun runa

DEM big man

'that big man'

It seems, therefore, that we need a rather sophisticated approach to lexical word classes if we want to take into consideration facts from the many less familiar (i.e. non-European) languages that are spoken across the globe. Such an approach is provided in Hengeveld (1992a, b), whose classification of parts-of-speech systems is discussed below.

1.5.3. Hengeveld's classification of parts-of-speech systems

In Hengeveld's view two basic types of language should be distinguished: (i) flexible languages, which have at least one major class of multifunctional lexemes (V/N/A or N/A), and (ii) rigid languages, which only have one or more major classes of specialized lexemes (V-N-A, V-N, or V). It must be added that Hengeveld (1992b: 58) states that 'languages at best show a strong tendency towards one of the types'.

Hengeveld's classification actually has seven types, because he also includes (manner) adverbs. However, since adverbs are irrelevant in the present context I have collapsed three of Hengeveld's types (V-N-A/adv, V-N-A-adv, V-N-A) into one (type 3: V-N-A) in Fig. 1.1. In defining the word classes Hengeveld took as his starting point the function of a predicate in a linguistic expression. Four major functions are usually distinguished: (1) head of the clause (verbal function), (2) modifier of the head of the clause (adverbial function; note that Hengeveld only refers to manner adverbs in his classification), (3) head of the term or NP (nominal function), and (4) modifier of the head of the term (adjectival function).

	Type 1	V/N/A		
Flexible	Type 2	V	N/A	
	Type 3	V	N	A
Rigid	Type 4	V	N	—
	Type 5	V	—	—

Fig. 1.1. Parts-of-speech systems (based on Hengeveld 1992b)

(p.17) He uses the following definitions (1992b: 58):

A *verbal* predicate is a predicate which, without further measures being taken, has a predicative use *only*.

A *nominal* predicate is a predicate which, without further measures being taken, can be used as the head of a term (NP).

An *adjectival* predicate is a predicate which, without further measures being taken, can be used as a modifier of a nominal head.

An *adverbial* predicate is a predicate which, without further measures being taken, can be used as a modifier of a non-nominal head.

In certain languages some or all of the functions mentioned above are clearly distributed over distinct, non-overlapping groups of predicates (specialized predicates; types 3–5); in other languages some or all of these functions can be performed by the same group of predicates (flexible predicates; types 1–2). Languages of type 1 (Samoan), 2 (Quechua), and 5 (or rather type 4/5: Cayuga) have already been discussed in 1.5.2 above, so I will only give examples of types 3 and 4.

The Australian language Kayardild belongs to type 3, because it has one class of words that can only be used as verbs, another distinct word class that can immediately fill the head position in the NP (i.e. nouns), and a major class of about one hundred adjectives, i.e. words that ‘can normally only appear when qualifying an overt [NP] head. Thus in normal contexts *jungarra dalija* [big came] is unacceptable, as is *dathina jungarra dalija* [that big came]; an entity nominal like *dangkaa* “person” is necessary, as in (*dathina*) *jungarra dangkaa dalija* “(that) big man came” (Evans 1995: 234; cf. also p. 238).

Galela, finally, is a clear example of language without a distinct class of adjectives. If we take the Galela equivalent of the English adjective ‘big’ *lamo* and let it function as a modifier of the noun, we must also add a third person pronoun. This is because in Galela this property ‘(be) big’ is expressed through a verbal predicate whose sole argument must be explicitly expressed in the form of a pronominal element. Thus the whole phrase is rather like a subordinate clause. Furthermore, if used attributively, the first syllable of the verbal predicate in question is reduplicated, yielding the participial form.

Galela (van Baarda 1908: 35)

(12) awi °dòhu i lalamo
his foot it big:PRT
‘his big foot’

In Table 1.3 all the sample languages are classified in terms of Hengeveld’s parts-of-speech systems:¹⁷ Languages of intermediate types 3/4 resist straightforward (p.18)

Table 1.3. Parts-of-speech systems of languages in the sample

Type 1	V/N/A	Samoan
Type 2	V- N/A	Human, Imbabura Quechua, Turkish
Type 3	V-N- A	Abkhaz, Alambalak, Basque, Berbice Dutch Creole, Bukiyp, Burushaski, Dutch, Georgian, Guaraní, Hittite, Hmong Njua, Hungarian, Ika, Kayardild, Ket, Nama Hottentot, Nasioi, Ngalakan, Ngiti, Sumerian, Wambon
Type 3/4	V-N(- A)	Babungo, Bambara, Chukchi, Gude, Kisi, Oromo, Pipil, Sarcee, Tamil
Type 4	V-N	Burmese, Galela, Hixkaryana, Koasati, Korean, Krongo, Lango, Mandarin Chinese, Nivkh, Nung, Nunggubuyu, Tsou, Vietnamese, West Greenlandic
Type 4/5	V(-N)	Cayuga
Type unknown		Etruscan, Meroitic, Nahali

classification in that they only have a smallish, closed class of adjectives. Consider the following remarks on Tamil adjectives (Asher 1982: 186–7):¹⁸

The question of whether it is appropriate to recognise a separate morphological category of adjective in Tamil has long been debated, on the grounds that all but a very small handful of adjectival modifiers of nouns are derived forms. The set of those that cannot by simple rules be derived from noun or verb roots comprises such high-frequency items as *nalla* ‘good’, *periya* ‘big’, *cinna* ‘small’, *putu* ‘new’, *pazaya* ‘old’, and a few basic colour terms...

Most other adjectives are either derived from verb roots (e.g. *kette* ‘bad’, which in morphological terms is the past relative participle of *ketu* ‘get spoiled’) or formed by the addition of one of two adjectivalising suffixes.

Adjectives will be discussed in greater detail in section 4.3. For reasons discussed above Cayuga has been classified as intermediate type 4/5 (section 1.5.1).

(p.19) 1.5.4. Noun phrases¹⁹

Generally speaking languages use either pronouns (such as *you* and *she*) or lexical NPs (like *the man*, *two cars*, *those big trees*) to refer to entities. In this book I will mainly restrict myself to lexical NPs which are headed by an underived noun and which are used to refer to a single, discrete, first order (i.e. spatial) entity or to a single *individual*²⁰ Thus I will generally ignore proper names (*Max*, *Hanna*), NPs headed by derived nouns (e.g. *doormat*, *bus driver*), as well as NPs referring to

- collectives (e.g. *the family*);
- non-discrete entities, such as masses (e.g. *the gold*);
- abstracts (e.g. *sorrow*) and other non-spatial entities, such as *the wedding*, which refers a second order (or temporal) entity, or *the opinion*, which refers to a third order entity (a propositional content, which is true or false, and which can be asserted or denied, remembered or forgotten; cf. Lyons 1977: 442 f. and Dik 1997: 292).

I will also ignore NPs that require a special interpretation with respect to the number of individuals that are involved, e.g. NPs with generic reference such as ‘the whale’ in ‘The whale is a mammal’ or ‘Whales are mammals’. Note incidentally that generic noun phrases are also very restricted with respect to the kind of adnominal modifiers they can accommodate.

Above we saw that not every language is deemed to have a class of nouns. Now we will see that even if a language has nouns, whether flexible or specialized, this does not necessarily mean that it also has hierarchically structured phrases with a head noun and fully integrated attributive modifiers as in e.g. English ‘this_{Dem} big_A dog_N’.

1.5.4.1. Integral and non-integral noun phrases

It has often been observed that there are languages in which noun modifiers (if we can still call them that) are not fully integrated constituents of the noun phrase. For instance, Blake (1983: 145) argued that in the Australian language Kalkatungu ‘there are in fact no noun phrases, but ... where an argument is represented by more than one

word we have nominals in parallel or in apposition ... Each word is a constituent of the clause' (more on apposition in section 1.5.4.2 below).

(p.20) Kalkatungu (Blake 1983: 145):

(13)	a.	Cipa-yi	ጀtuku-yu	yaun-tu	yani	icayi
		this-ERG	dog-ERG	big-ERG	white_man	bite
	b.	Cipayi	ጀtukuyu	yani	icayi	yauntu
		this:ERG	dog:ERG	white_man	bite	big:ERG
	c.	ጀtukuyu	cipayi	icayi	yani	yauntu
		dog:ERG	this:ERG	bite	white_man	big:ERG
	d.	Yauntu	cipayi	ጀtukuyu	icayi	yani
		big:ERG	this:ERG	dog:ERG	bite	white_man
	e.	Cipayi	icayi	yani	ጀtukuyu	yauntu
		this:ERG	bite	white_man	"dog:ERG	big:ERG
	f.	Yani	icayi	cipayi	yauntu	ጀtukuyu
		white_man	bite	this:ERG	big:ERG	dog:ERG

'This big dog bit/bites the white man'

Similar observations have been made in connection with other Australian languages such as Mangarayi (Merlan 1982: 49), Ngalakan (Merlan 1983: 83), Ngiyambaa (Donaldson 1980: 232), Nunggubuyu (Heath 1984: 505–6), Wardaman (Merlan et al. 1997: 86–7), and Gooniyandi (McGregor 1989, 1997b), where the same phenomenon is also discussed under the headings of *configurational* vs. *non-configurational*, *tight* vs. *scrambled*, or *hierarchical* vs. *flat*.²¹

In some languages appositional modification is related to complexity in the prefield of the noun. If the number of prenominal modifiers exceeds a certain limit, one modifier will typically occur after the noun in an appositional relationship.²² This happens, for instance, in Yimas of the Indo-Pacific phylum. In this language (Foley 1991: 4, 184, 188)

[NPs] may consist of only two constituents, a modifier and a head, in that order. If more than two modifiers are present for a given noun, one must take an agreement suffix and occur in the scrambled pattern. ... The scrambled structures have very different properties from the tight noun phrases. They are not single noun phrases at all, but rather two noun phrases in apposition, one consisting of a noun, the other a modifier, nominalized by the agreement suffix. They are joined together as a single semantic unit (but two syntactic units) by the agreement suffix, for it is this marker on the modifier which links it to the noun being qualified.

(p.21) In other languages only certain noun modifiers are regularly apposed, notably numeral classifier phrases that follow the noun (or better: the term containing the head noun). For instance, Korean has two alternatives when it comes to the placement of numeral + classifier: the integrated prenominal variant (14b), and the more popular

postnominal variant (14a), which ‘may best be described as a special type of appositional construction’ (Lee 1989: 118).²³

Korean (Lee 1989: 118)

(14) a.	cɛg	isibo	gwʌn
	book	twenty-five	volume
‘twenty-five books’			
b.	isibo	gwʌn-ii	cɛg
	twenty-five	volume-II	book
‘twenty-five books’			

The following example from Nama Hottentot ‘can be considered as consisting three miniature noun phrases in apposition, each marked as third person femin plural by the postposition of *-di* and each capable of occurring in isolation.’ flavour of the construction is captured by an English version like: *books, th three, the ones of !Gombate*. The whole construction, however, forms a sin phrase to which a case-assigning postposition may be attached’ (Payne 1994: 2850; cf. also Hagman 1974: 90 f.).

Nama Hottentot (Payne 1994: 2850)

(15)	#kini-di	ne	!nona-di	!Gombates di-di
	book-3PL.F	these	three-3PL.F	!Gombates of-3PL.F
‘these three books of !Gombate’s’				

It seems that in cases such as these we do not have apposition at the sentence level, as in the Kalkatungu example above, but rather apposition of minor NPs denoting the same referent within a larger phrasal structure that is marked for case.

(16) ([books]_{NP} [these three]_{NP} [the ones of !Gombate]_{NP})_{NP}

As a matter of fact modifiers in Nama can also appear as fully integrated constituents, but then they always precede the noun and are not marked for person, number, and gender.²⁴ Similar cases of asymmetry are found in e.g. Georgian and in Samoan. In Georgian ‘a postposed modifier behaves more like a separate (p.22) NP, or apposition in that it inflects for the full set of cases and numbers’ (Testelec 1991b: 652, 677):

Georgian (Testelec 1991b: 652)

(17) a.	am	or	lamaz	kal-s
	that:OBL	two	nice	woman-DAT
‘to those two nice women’				
b.	kal-eb-s		lamaz-eb-s	
	woman-PL-DAT		nice-PL-DAT	

'to the nice women'

In Samoan postnominal demonstratives that combine with an article are regarded as appositional modifiers (more on this in Chapter 6):

When following the noun phrase nucleus, the forms *lēnei*, *lenā*, *lelā*, and *lea* constitute a noun phrase consisting of the article and the demonstrative nucleus, which is syntactically classified as an apposition. A literal translation of '*o le tama lea*' would be 'the boy, that one'. Inherently specific noun phrases are determined by demonstratives which without an article directly follow the nucleus. (Mosel and Hovdhaugen 1992: 292)

It may also be the case that the equivalent of a fully integrated noun modifier in one language must be regarded as a separate constituent at the level of the sentence in another language. This appears to hold for the Hixkaryana equivalent of English attributive numerals, which according to Derbyshire (1979: 103) are basically adverbs.

Hixkaryana (Derbyshire 1979: 44)

(18) a. asak kanawa wenyo	b. kanawa wenyo, asako
two canoe l:saw:it	canoe l:saw:it two
'I saw two canoes'	'I saw two canoes'

1.5.4.2. Apposition

In the previous section we saw that several languages are stated to have noun modifiers that are actually part of an appositional construction. However, the term *apposition* stands for a variety of construction types in the linguistic literature (cf. Matthews 1981: 224–36) and in grammatical descriptions of individual languages this term usually remains undefined.

I use the term *appositional modifier* for all those elements which semantically speaking serve the same purpose as their non-apposed, integrated counterparts, but which from a syntactical point of view are not part of the (integral) phrase containing the head noun. An obvious requirement of any apposed modifier phrase is co-reference: it must refer to the same entity as the other member(s) in the appositional construction.²⁵ In the English paraphrase of an appositional (p.23) construction this is typically indicated by the employment of the dummy head 'one', as in this example from Herring (1991: 55), in her article on attribution in Tibeto-Burman languages (see also the example from Nama Hottentot in section 1.5.4.1 above):

Relative and attributive clauses ... are not adjectival ... but rather function in some sense as nominal appositives (e.g. 'the [banana-eating one] boy'; 'the [fat one] boy').

This paraphrase also illustrates that syntactically speaking an appositive modifier can in principle serve as an independently referring constituent by itself (which would at least partly explain why appositives may occur in a discontinuous sequence in the clause—as in the Kalkatungu example above). However, since appositive and integrated noun modifiers do not differ in their semantic (qualifying, quantifying, localizing) function, an appositive modifier usually appears together with the 'head' noun (which of course

cannot be regarded as the *syntactic* head of the appositional modifier—hence the quotation marks), just like a modifier that is part of an integral NP.

Although it is often unclear how the term ‘apposition’ should be interpreted in the various language descriptions, all noun modifiers that are stated in these grammars to be in an appositional relationship with the noun seem to share one property: they are not regarded as immediate dependants of the noun, i.e. appositional modifiers are not in a tight head-dependent relationship that is characteristic for the constituents that make up an integral noun phrase. Or, in terms of constituency, an appositive is not deemed to be part of a hierarchically organized constituency structure with the noun as its semantic and syntactic nucleus. Since the main topic of this study is the structure of the *integral* NP, I will exclude instances of appositive modification from the current investigation.²⁶

1.5.4.3. Complexity

Both within and across languages NPs may vary considerably with respect to their internal organization and complexity.²⁷ It appears, for instance, that in written language NPs tend to be more complex grammatically than NPs used in spoken language (Hindle 1981; Linell 1982; Perkins 1992: 89). Furthermore, factors such as politeness, taboo situations, and familiarity may co-determine the internal composition of linguistic expressions (Haiman 1985a; see also below). (p.24) Thus, an NP may involve just a noun as in

(19) *Houses* are expensive these days

but it may also consist of a large string of elements, as in

(20) I would like you to meet *the two brave men who saved my life last year when I rather foolishly decided to climb that mountain in Nepal from which I had to be rescued because I had lost all my gear in a blizzard*.

Such differences in complexity are mostly due to such considerations as are captured in Grice’s maxims of conversation, notably his maxim of quantity, which reads (Grice 1975: 45–6):

- (i) make your contribution as informative as is required for the current purpose of the exchange;
- (ii) do not make your contribution more informative than is required.

Although it is in theory possible to build extremely long and complex NPs, it is stated in many grammars that in normal, unelicited speech an NP will most frequently involve only two, maximally three, constituents. This is usually attributed to stylistic or processing factors.²⁸

It may be useful to distinguish between phrasal and non-phrasal modifiers when we speak of the internal complexity of NPs (cf. also Dryer 1991). It has been observed that recursively embedded phrasal modifiers such as possessor phrases and relative clauses are attested relatively more frequently than multiple non-phrasal lexical modifiers, i.e. adjectives. This may have to do with the fact that possessor phrases and relative clauses are typically used for identifying purposes, i.e. they make it possible for the hearer to correctly identify the referent of the matrix NP (Chapter 6). The following example is from Babungo and involves a four-step identification procedure. In this sentence *tò* ‘head’ ('hair') is identified through ‘Lambi’, then ‘Lambi’s wife’, then ‘Lambi’s wife’s brother’, and finally ‘the child of Lambi’s wife’s brother’:²⁹

Babungo (Schaub 1985: 76)

(21) ḥwə́ bá	tò	wéé	wéenshú	zú	Làmbí
she	weave:PRES	head	child	brother	wife Lambi
'(she is plaiting) the hair of the child of Lambi's wife's brother'					

It has also been observed that in some languages speakers prefer to distribute many modifiers over several NPs rather than have them all in one NP, as in Pacöh (Watson 1976: 220): ‘A fully expanded noun phrase is not expected. Rather, the (p.25) speaker tends to give the information in two or three short phrases or clauses’ (see also e.g. Miller 1964: 65 on Brôu). Alternatively, in some languages modifiers must be coordinated so as to avoid excessively complex structures, as in Katu: ‘It is preferred to break a long phrase, using two or more co-ordinate phrases instead’ (Costello 1969: 22). Interestingly a similar phenomenon occurs in relation to the number of NPs in a clause. It has been reported that speakers of certain languages tend to distribute (many) NPs over more than one clause, so as to avoid a clause having too many NPs.³⁰

There is probably also a correlation between the pragmatic or discourse function of an NP in a clause on the one hand and its internal complexity on the other, which may be summarized as follows: what is familiar is given reduced expression (Haiman 1985a: 18, referring to Zipf 1935; see also below). Thus, NPs referring to established discourse referents will usually not contain much linguistic material; often a pronominal element will suffice for successful communication (cf. Grice’s maxim of quantity above). This principle has been given different formulations. Consider, for instance, Givón’s code-quantity principle, according to which NPs denoting topical or ‘known’ referents will usually not contain many constituents (Givón 1988: 249):

the less predictable/accessible a referent is, the more phonological material will be used to code it.

Thus NPs referring to new referents tend to contain relatively much linguistic material, especially if the referents of these NPs are presented as identifiable, i.e. definite (like ‘the two brave men’ in (20) above; cf. Rijkhoff 1989).³¹

Culture also plays a role with respect to NP-internal complexity. Perkins (1992: 24, 191) found a positive correlation between cultural complexity and the number of NP positions that can be relativized as well as a negative correlation between cultural complexity and the number of grammatical deictic distinctions. Haiman (1985a: 149–50) has argued that languages tend to have complex periphrastic means of expressing notions that are unfamiliar, which he calls the ‘iron horse’ (= locomotive) phenomenon; the illustrative examples are from Hmong and Nivkh.³²

Hmong (Downing and Fuller 1984)

(22) daim ntawv qhia ev boor
CLF paper tell way mountain
‘map’

(p.26) Nivkh (Mattissen and Drossard 1998: 4; original example in Jakobson 1971a: 80)

(23)	pəi-mu-meñ-vo-ñivx
	fly-boat-helm-take-person
	'aviator'

Style is yet another factor that co-determines the degree of complexity in the NP. Haiman (1985a: 151 f.) has pointed to the fact that polite speech requires longer, and consequently more complex, linguistic expressions. For instance, the most polite way to ask in Javanese 'Are you going to eat rice and cassava now?' would involve twenty-two syllables, whereas the most familiar version has only fifteen syllables (see also Dik 1986: 27–8).

Javanese (Haiman 1985a: 151)

(24) a.	menapa	nandalem	mundhut	sekul	semanten
b.	menapa	panjenengan	mendhet	sekul	semanten
c.	napa	sampeyan	mendhet	sekul	semonten
d.	napa	sampeyan	njupuk	sega	semonten
e.	apa	sliramu	mundhut	sega	semono
f.	apa	kowe	njupuk	sega	semono
	QM	you	take	rice	that_much

'Did you take that much rice?' (QM = question marker)

It has also been established that in some languages there is a correlation between the internal composition of the NP (in terms of choice and variation of linguistic elements) and the sex of the speaker. For instance, in Hmong each regular classifier has different forms, which correlate with the speaker's evaluation of a particular referent: (a) grand and imposing, (b) ordinary, (c) minute. Women, however, tend to use only one form of the classifier (Wang 1972: 160).³³

Finally, Capell (1965: 452) distinguished between *event-dominated* and *object-dominated* languages (among others). Object-dominated languages have a rather rich morphology and syntax in the NP (e.g. noun class system, extensive agreement) and a relatively simple verb. Conversely, event-dominated languages (p.27) typically have a rich verbal system, whereas the internal syntax and morphology of NPs is relatively simple (cf. also de Vries 1989: 77).

In sum, there is wide variety of factors that co-determine the internal complexity of NPs. In this book, however, I will mainly be concerned with the following questions: what kind of nouns and noun modifiers are attested in the languages of the world (Chapters 2–6), what are their morpho-syntactic properties (Chapters 8–10), and what does a typologically adequate underlying structure of the integral noun phrase look like (Chapter 7)?

1.5.5. Referents of NPs

Noun phrases are referring expressions, but the entities they refer to are not entities in the external physical world. Referents of NPs are rather mental representations of entities as they are created, stored, and retrieved in the minds of the speech participants. This can be demonstrated by the fact that we can talk about entities that are not actually present in the speech situation (the Great Wall, the man who killed John Lennon), entities that never existed (unicorns, the monster of Loch Ness, Martians) or fictional characters (Odysseus, Don Quixote, Tom Sawyer, Molly Bloom).³⁴ Thus, a distinction must be made between a referent (a mental construct) and its ontological counterpart in the physical world, if it exists, for not every referent has such a ‘Sein-correlate’ (Rapaport 1985/6). Since referents of NPs are mental constructs that are construed on the basis of linguistic material rather than real-world objects, we will see below in Chapter 2 that this allows for the possibility that there are certain (aspectual) discrepancies between properties of referents and properties of their real-world counterparts.³⁵

Notes:

- (1) Cf. Sanders (1976: 161): ‘Like all other instrumental objects ... it is the function of grammars that determines both their essential characteristics and their range of possible variation and change.... Any theory that fails to generate ... functionally based explanations will simply fail to be a real theory about grammars, in precisely the same sense that a theory referring only to the physical properties of hammers, or the psychology of the hammer-users, would fail to be a real theory about hammers. An instrumental object, in other words, *exists* in a very real sense only because of its function, and can be significantly investigated and understood, therefore, only in terms of that function.’
- (2) There are quite a few language specific studies about NPs—see for example Börjars (1994) on Swedish; Costello (1969) on Katu; Delsing (1993) on all Scandinavian languages; Gangopadhyay (1990) on Bengali; Gao (1994) on Chinese; Holm (1990) on the Atlantic Creoles; Kölver (1978) on Newari; Szabolcsi (1994), Lacko (1995), and Moravcsik (1994,1997) on Hungarian; Miller (1964) on Brôu; Schroeder (1999) on Turkish; Thomas (1976) on Chrau; Watson (1976) on Pacöh; Willim (1995) on Polish. See also the contributions in Siewierska (1991b) and the forthcoming EuroTyp volume on noun phrase structure in the European language (Plank forthcoming). By now there are also quite a few grammatical descriptions of languages according to questionnaires that also include detailed questions about NPs, such as the volumes of the Lingua Descriptive Studies (first published by North-Holland, now by Croom Helm) and the descriptions in the *Handbook of Amazonian languages* (Derbyshire and Pullum 1986, 1990, 1991, 1998). Some general (both theoretical and empirical) studies about the structure of NPs are: Abney (1987); Bach (1968); Benveniste (1966); Dik (1997: 127–91); Ewert and Hansen (1993); Foley (1980); Gil (1987); Hawkins (1994); Heine (1980); Jones (1970); Lappin (1988); McCawley (1971); Meillet (1906); Payne (1994); Rijkhoff (1990a, b); Seiler (1978a, 1985); Valois (1991); van der Auwera (1990); Van Valin and La Polla (1997: 52–69); Zimmermann (1991).

- (3) Preliminary reports on these topics appeared in *Working Papers in Functional Grammar*, 29 (= Rijkhoff 1988), *Linguistics*, 28/1 (= Rijkhoff 1990a), Nuysts et al. 1990 (= Rijkhoff 1990b), *Journal of Semantics*, 8/4 (= Rijkhoff 1991), *Gramma/TTT*

tijdschrift voor taalwetenschap, 4/3 (= Rijkhoff 1995), and Vogel and Comrie 2000 (= Rijkhoff 2000).

(4) The current sample is based on the classification of languages proposed in Ruhlen (1987) and has been modified according to the changes that Ruhlen proposed in the postscript of the 1991 edition, in which he writes that '[t]he family called Caucasian ... should now be broken up into two distinct families: North Caucasian and South Caucasian ... Following a suggestion by Sergei Starostin I now use Caucasian in place of North Caucasian, and Kartvelian in place of South Caucasian. Furthermore, I would now remove Korean-Japanese-Ainu from Altaic, leaving Altaic with its traditional constituents: Turkic, Mongolian, and Tungus' (Ruhlen 1991: 379). Notice that the non-genetic family of thirty-eight Pidgins and Creole languages in Ruhlen's classification is treated the same way as the other families in this sampling procedure.

(5) On Nahali see Grierson (1906: 185–9) and Kuiper (1962); on Etruscan see Stoltenberg (1950) and Bonfante (1983); on Meroitic see Haycock (1978) and Campbell (1991: 920–2).

(6) See Andrews (1985) and Dik (1997) for semantic, pragmatic, and syntactic (grammatical) functions of NPs.

(7) See, for example, Lyons (1977: chapter 11); Backhouse (1984); Plank (1984); Schachter (1985); Thompson (1988); Pustet (1989, 2000); Croft (2000); Wetzer (1992); Sasse (1993); Bhat (1994); Anward et al. (1997); Beck (2000); Prasithrathsint (2000).

(8) Although a few R2 roots tend to occur in more or less lexicalized forms, they can still be used as the head of the clause, e.g. *kaqtanéhkwi* ‘it pulls logs, horse’, *teká:téh* ‘it habitually goes up, airplane’; *qtwenotákhwa* ‘one habitually puts one’s voice in it, telephone’ (Sasse 1993: 207).

(9) Cf. Sasse (1993: 203; also 1988: 186 ff.): ‘Im Cayuga sind alle in aktuellen Äußerungen erscheinenden Inhaltswortformen syntaktisch prädikativ, d.h. ohne weitere Hilfsmittel geeignet zum Ausdruck einer eigenständigen, vollständigen Proposition. Sie repräsentieren damit eine Äußerung, die in europäischen Sprachen Satzcharakter hätte’ [In Cayuga all content words that appear in actual utterances are syntactically predicative, i.e. no further measures are required to express an independent, complete proposition. Thus they represent an expression that would constitute a sentence in European languages].

(10) See also Mithun (1996: 634, 636) on possession in Mohawk: ‘Possessive constructions appear less often in Mohawk than in many languages because of the existence of alternative structures that accomplish some of the same functions. Kinship, for example, is generally expressed by means of stative verbs that relate participants A second reason for the comparative rarity of nominal possessive constructions in Mohawk is the fact that many entities are identified by means of morphological verbs rather than nouns.’ Cf. also Bonvillain (1973: 25, 221, 199) on Akwesasne Mohawk.

(11) In Chapter 7 of this book I will draw parallels between the underlying structures of the NP and another linguistic unit (namely, clauses, or rather predication, as defined in FG; see Dik 1997).

(12) See on Samoan word classes also Churchward (1951: 126; as cited in Vonen 1994: 155): ‘Almost any part of speech can be used as any other part of speech.’

(13) The text continues as follows: ‘Not all roots occur with the same frequency as verbs and nouns. Some roots predominantly function as verbs, whereas others are more likely to be found in the function of nouns. Until now we have not, for instance, found *alu* “go” in a nominal function or *mea* “thing” in a verbal function ... But we hesitate to say that *alu* is inherently a verb and *mea* inherently a noun for two reasons. Firstly, we cannot find any functional explanation why *alu* should not be used as a noun and *mea* as a verb, whereas, for instance, *gaoi* “thief, to steal” and *tagata* “person, to be a person” are bi-functional. And, secondly, previous experience taught us to be careful with classifications. The more texts we analyzed, and included in our corpus, the more items were unexpectedly found in nominal or verbal function.’

(14) There are, however, a number of derived predicates carrying the causative prefix *fa* ‘*a-*, which typically occur as modifiers of the head of the term phrase (Mosel and Hovdhaugen 1992: 73, 119, 175).

(15) See also, for example, Jelinek (1995) and Jelinek and Demers (1994) on Straits Salish. The qualification ‘multifunctional’ has been taken from Broschart (1991).

(16) Cf. also Churchward (1953: 16) on Tongan: ‘In Tongan ... there is much interchange of functions between the various parts of speech. This applies particularly to nouns, verbs, adjectives, and adverbs.’

(17) See also Hengeveld et al. (forthcoming). The fact that parts-of-speech systems are not evenly distributed among the languages in the sample probably reflects global distribution patterns. It is not difficult, however, to find other languages of types 1, 2, and 5. For example, languages with one major flexible word class (type 1: V/N/A) are not only attested in other Polynesian languages such as Tongan (Broschart 1991, 1997; Churchward 1953: 16), but also in e.g. Mundari, which is an Austro-Asiatic language (Hoffmann 1903: pp. xx–xxi; Sinha 1975: 76), and many of the Turkic languages do not distinguish between nouns and adjectives (type 2: V-N/A; see, for example, the contributions in Deny et al. 1959).

(18) On adjectives as a minor word class, see also Brauner (1974: 38) on Bambara; Hoskison (1983: 53) on Gude; Childs (1995: 126) on Kisi; Lee (1989: 40) on Korean; and Campbell (1985: 120) on Pipil. The situation in Sarcee, however, is not entirely clear. Although the two most frequently used qualifiers (*tsítł'á* and *tcúw*) ‘are obviously related to verb stems *-tsítł'á* “to be small” and *-tców* “to be big”, Cook (1984: 67) also writes that not every qualifier is traceable to a verb stem.

(19) Often ‘noun phrase’ (NP) and ‘term (phrase)’ are used interchangeably, but strictly speaking ‘term’ or ‘term phrase’ has a wider application than ‘noun phrase’. The latter only involves phrases headed by a noun, whereas ‘term (phrase)’ may also apply to e.g. object clauses, headless constructions, or phrases consisting of pro-nouns. Of late ‘term’ has even been employed for any referring expression (Keizer 1991). Originally, in its logical sense, terms (Latin *terminus*) were atomic elements that had no structure of their own (cf. Vendler 1971: 116; Lyons 1977: 148).

(20) On nominalization see Lehmann (1982a); Comrie and Thompson (1985); Mackenzie (1987a); Koptjevska-Tamm (1993).

(21) McGregor (1989: 219), however, distinguishes between *discontinuous* and *fractured* NPs. Phrasal discontinuity (which is a rather typical characteristic feature of many Australian Aboriginal languages) is ‘principally a means of allowing a single nominal phrase to fulfil two distinct textual roles—theme and unmarked focus ... , or marked focus and elaborating adjunct’. Fracturing, on the other hand, ‘may be characterized as a means of permitting each constituent of a nominal phrase to fill one and the same textual role’. He adds that fractured NPs are definite, whereas discontinuous phrases are typically indefinite.

(22) Postnominal placement due to complexity in the prenominal area does not seem to result in forms of appositional modification in e.g. Bukiyp (Conrad 1991: 57) and Wambon (de Vries 1989: 77).

(23) On the appositional nature of the numeral classifier phrase, see also e.g. Jones (1970: 7); Hope (1974: 89) on Lisu; Greenberg (1972: 28, 1975); Goral (1978: 12, 21); Comrie (1981a: 269) on Nivkh; Lehmann (1982b: 255); Foley (1986: 78); Payne (1986: 113); Court (1987: 145 f.) on Yao; Wheatley (1987: 851) on Burmese; Derbyshire and Payne (1990: 243); Corbett (1991: 311); see also Croft (1990a: 31) on possessive classifiers.

(24) Cf. also Vilkuna (1997: 226) on postnominal genitives in Erzya (an Uralic language), which are said to be ‘more loosely connected’ to the head than the prenominal ones.

(25) This also holds for clause-level appositives. Mithun has argued that in e.g. Cayuga argument NPs function as appositives to the pronominal affixes in the verbal complex (Mithun 1992: 59): ‘In pragmatically ordered languages, separate noun phrases can function somewhat differently than in languages without bound pronouns. They typically serve more as appositives to the bound pronouns than as primary arguments themselves.’

(26) See McGregor (1997a: chapter 3) for an overview of the various kinds of grammatical relations that may hold between constituents.

(27) Differences in linguistic or grammatical complexity have been discussed in the context of phonology/morphology (Haugen 1976: 286; Braunmüller 1995: 77; Werner 1984: 220), morphology/syntax (Dik 1997: 392), lexicon (Aronoff 1994), number of speakers (Whinnom 1980: 212; Hymes 1971: 73; Mühlhäusler 1996; Trudgill 1992: 204), culture/deixis (Perkins 1992), language shift and language contact (O’Neill 1982: 285; Thomason and Kaufman 1988), language acquisition and grammaticalization (Labov 1990; Bickerton 1981; Trudgill 1992). I am grateful to Wouter Kusters and Hadewych van Rheeden, to whom most of these references are due.

(28) See Bruce (1984: 100) on Alambalak; Schaub (1985: 76) on Babungo; Dixon (1988: 36, 117) on Boumaa Fijian; Gregores and Suárez (1967: 150) on Guaraní; Testelec (1997a: 236) on Georgian; Lee (1989: 120) on Korean; Reh (1985: 252) on Krongo.

(29) See also Saltarelli (1988: 80) on Basque; Lee (1989: 120) on Korean; Rausch (1912: 120) on Nasioi; Thomsen (1984: 54) on Sumerian; Asher (1982: 63) on Tamil; Fortescue (1984: 144) on West Greenlandic.

(30) See e.g. Derbyshire (1979: 39) on Hixkaryana; Gorbet (1976: 155) on Diegueño; Davies (1981: 45–6) on Kobon; Reesink (1984: 133–4) on Usan.

(31) Consider in this context also this observation by Watson (1976: 220) on the internal complexity of subject NPs, which tend to refer to topical referents: '[In Pacōh] noun phrases filling the subject slot of the clause are usually simple with little modification. The most fully expanded noun phrases are found in the object slot of the clause.'

(32) See also Foley (1997: 36–7) on similar phenomena regarding verbs. Compare, for example, the Kalam (Indo-Pacific) translational equivalent of English ‘I fetched some firewood’ (from Pawley 1993: 95): *yad am mon p-wk d ap ay-p-yn* [1SG go wood hit-break hold come put-PERF-ISG]).

(33) It is easier to find examples outside the NP-internal domain. For example, in Samoan ‘women use subject-initial word order far more often than men (four times as much overall)’ and ‘men use verb-subject-object word order more than women (nearly twice as much overall)’. On the other hand, differences in the use of ergative case marking also depend on features of the addressee: ‘For ergative case marking, it was found that men use this marking as often as women do in family interactions but much more often than women in interactions involving nonfamily members’ (Ochs 1987: 66). For a cross-linguistic overview of men’s and women’s speech, see Bodine (1975). For other examples of differences between male and female speech, see e.g. Philips et al. (1987); Levinson (1988); Foley (1997: 299f.); Teferra (1987); especially Ide and Hill (1998); also Bruce (1984: 320) on Alamblak; Comrie (1981a: 244) and Dunn (2000) on Chukchi; Kimball (1991: 53, 224; see also Kimball 1987) and Haas (1944) on Koasati; Lee (1989: 72) on Korean; Thomsen (1984: 285 ff.) on Sumerian; cf. finally also Mithun (1999: 276–80) on ‘men’s and women’s language’ and other special speech styles in the languages of native North America, and Suárez (1983: 55) on the use of whistled speech in Mesoamerican languages.

(34) Hence we speak of the *intended* referent of an NP, rather than its *actual* referent (McCawley 1968: 138; Dik 1997: 127).

(35) Thus ‘referents’ should not be interpreted in the metaphysical sense, but rather as entities that figure in the world of discourse, i.e. as ‘discourse referents’ (Karttunen 1976).

2 Nominal Subcategories: Seinsarten

2.1. Introduction

In Chapter 1 it was established that most languages either have a distinct class of nouns (N) or that they use what we may call flexible nouns (V/N/A, N/A), i.e. lexemes that occur as the head of the noun phrase (the prototypical function of a noun) but which are also used as the head of the sentence and/or to modify the head of the term ('without further measures being taken'—cf. Hengeveld 1992b: 58). In this chapter I will investigate certain morpho-syntactic and semantic properties of both flexible and distinct nouns that are employed in languages across the world to refer to a singular discrete spatial entity (or *individual*) such as 'dog' or 'knife' (or, rather, to refer to an entity whose ontological counterpart or *Sein*-correlate is a singular discrete spatial object).¹

I will argue on the basis of morpho-syntactic evidence that four noun types (or nominal subcategories) are normally used for this purpose: *singular object nouns*, *set nouns*, *sort nouns*, and *general nouns*. Ultimately this leads to a more comprehensive and semantically orientated classification of six major nominal subcategories which are defined in terms of two features: Shape and Homogeneity (section 2.4). Section 2.5 is devoted to the semantics of incorporated and predicate nouns and in the last section I will contend that essentially each nominal subcategory defines a different *Seinsart* of a spatial property, i.e. an alternative way in which a nominal property is specified for the features Shape and Homogeneity (section 2.6).²

2.2. A preliminary classification

A cursory examination of nouns within and across languages reveals that first order nouns, i.e. nouns that are used for (discrete) spatial objects in the real world, do not all have the same semantic and morpho-syntactic properties. In this section (p.29) I will mostly be concerned with morpho-syntactic properties which have to do with quantification. That is to say, I will establish for all the languages in the sample

- whether or not first order nouns appear in the plural form when modified by a numeral (where $n > 1$);
- if first order nouns are in a direct construction with a numeral or whether the numeral must first combine with a classifier.

Logically speaking there are four possibilities (disregarding the differences in word and morpheme order):

1. numeral + noun + plural (no classifier)
2. numeral + noun (no plural, no classifier)
3. numeral + classifier + noun (no plural)
4. numeral + classifier + noun + plural

Previous research has already indicated that there are very few (possibly no) languages of type 4, i.e. languages in which the noun must take a plural marker while the attributive numeral combines with a true classifier (cf. Sanches and Slobin 1973; more

on this in Chapter 5, which also deals with the syntax of numerals). The other three types are exemplified below:

Dutch: numeral + noun + plural

(1) twee boek-en

two book-PL

'two books'

Oromo (Stroomer 1987: 59): numeral + noun

(2) gaala lamaani

camel two

'two camels'

Thai (Hundius and Kölver 1983: 172): numeral + classifier + noun

(3) rôm sám khan

umbrella three CLF:long, handled object

'three umbrellas'

Both in Dutch and in Oromo the numeral is in a direct construction with the noun (i.e. no classifier), but only in Dutch does the noun appear in the plural form. In Thai the numeral must combine with a numeral classifier and the noun remains unmarked for number. For reasons that will be explained in section 2.3 I will call nouns of the Dutch type *singular object nouns* and nouns of the Oromo type and the Thai type will be referred to as *set nouns* and *sort nouns*, respectively.

Table 2.1 gives an overview of the distribution of these noun types across the languages in the sample which have a *distinct* class of nouns. This excludes languages of types 1, 2, and 5 in Hengeveld's parts of speech typology discussed in Chapter 1. The sample contains a number of languages which defy (p.30)

Table 2.1. Nominal subcategories in sample languages: singular object nouns, set nouns, and sort nouns

	<u>num + N + Pl</u> singular object noun	<u>num + N</u> set noun	<u>num + Clf</u> sort noun
<u>Type 3: V-N-A</u>			
Abkhaz (see below)	+	+ ?(-human/+ animate)	
Alamblak (see below)			
Basque		+	
Berbice Dutch		+	
Bukiyip	+		
Burushaski	+	+	
Chukchi (see below)			

		+ (NUM = bound except if case = ABS)
Dutch	+	
Georgian		+
Guaraní (see below)		
Hittite	+ (N = Masc/Fem)	+ (N = Neuter)
Hungarian		+
Ika		+
Kayardild		+
Ket (see below)	+	+
Hmong Njua		+ (but see 4.3.6)
Nama Hottentot (see below)		
Nasioi	+ (N= + human)	+ (N = -human)
Ngalakan		+
Ngiti	+ (N= + human)	+ (N = -human)
Sumerian		+
Wambon		+
<u>Type 3/4: V-N(-A)</u>		
Babungo	+	
Bambara (see below)		
Gude		+
Kisi	+	
Oromo		+
Pipil	+ ?	+ ?
Sarcee	+ ?(N = + human)	+ (N = -human)
Tamil	+ (N = + human)	+ (N = -human)
<u>Type 4: V-N</u>		
Burmese		+
Galela		+
Hixkaryana		+ (see below)
Koasati (see below)		
Korean		+
Krongo	+	
Lango		+
Mandarin Chinese		+
Nivkh	+ (NUM = bound)	+ (but see below)

Nung		+
Nunggubuyu	+ (N= + human)	+
Tsou		+
Vietnamese		+
West Greenlandic	+	
<u>Type unknown</u>		
Nahali		+

(p.31) straightforward classification in terms of the criteria mentioned above. Apart from the fact that a language may employ more than one nominal subcategory (e.g. singular object nouns to refer to human or animate entities and set nouns to refer to non-human or inanimate entities), there are basically four kinds of problems. The first problem relates to the fact that in some languages number is not marked on the noun but on the noun phrase (e.g. Nama Hottentot). Then there are languages in which the numeral is not a noun modifier, but a sentence adverb (e.g. Hixkaryana), the head of a non-noun modifying clause (e.g. Koasati) or some other kind of non-attributive expression (such as the body-part counting systems of the Papuan languages, where counting involves touching the relevant body-part; see also Chapter 5). Thirdly, nouns may be in transition so that they display properties of ‘old’ subcategory X and at the same time exhibit features of ‘new’ subcategory Y (e.g. Nivkh). Finally, there are languages in which, depending on certain conditions (e.g. case, humanness), numerals are expressed as free modifiers or as morphologically bound forms (Abkhaz, Chukchi, and Nivkh). Some of these cases are briefly discussed below, but the same topics will also be addressed in more detail in Chapter 4 on Quality (where I will argue that the so-called number markers on set nouns are better analyzed as nominal aspect markers) and in Chapter 5 on Quantity (number marking and numerals).

Alamblak, Bambara, Guarani, and Nama Hottentot (but possibly also e.g. Quechua; see 2.2.4) are all languages in which it is the noun phrase rather than the noun that is marked for number.³ Both in Alamblak and Nama NPs are followed by a phrase-final enclitic element that indicates number, person, and, in the case of third person, also gender. This person-number-gender marker shows great similarity with a corresponding member of the set of personal pronouns (Bruce 1984: 96; Hagman 1974: 40 f., 84 f.). Third person markers are of course (p.32) most common; others forms, however, also occur in certain phrases ‘with an effect something like an appositional phrase in English’ (Bruce 1984: 96), e.g.

Alamblak (Bruce 1984: 96)

(4) a. yima-m	b. yima-r	c. yima-t
person-3PL	person-3SG.M	person-3SG.F
‘people’	‘man’	‘woman’
(lit. ‘they people’)		

(5) a. yima-kë	b. yima-nëm
person-2PL	person-1 PL

'you people' 'we people'

Nama Hottentot (Hagman 1974: 43)

(6) a. kxòe-`n	b. kxòe-p	c. kxòe-s
person-3PL	person-3SG.M	person-3SG.F
'people' ⁴	'the male person'	'the female person'

At least in Nama Hottentot the plural form seems compulsory, both with and without a modifying numeral:

Nama Hottentot (Hagman 1974: 68)

(7) !noná	kái	/íri-ku
three	large	jackal-3PL.M
'three large male jackals'		

In Bambara, too, the plural suffix (which is clearly related to the third person plural pronoun) is a phrase-final enclitical element, so that it is properly speaking not the noun but rather the NP that is marked for number (Brauner 1974: 26–8; Kastenholz 1989: 21). Despite the fact that the plural marker is absent when the noun is modified by a numeral, I have refrained from labeling Bambara nouns as set nouns.⁵

Bambara (Brauner 1974: 28)

(8) musa	fila
woman	two
'two women'	

The Guarani ‘qualifiers’ *kwéra* and *hikwái*, which are used to mark plural number (the former pluralizes nouns and pronouns, the latter verbs and, rarely, nouns), are also optional, phrase-final constituents.⁶ They appear after all other postnominal (p.33) noun modifiers and are mutually exclusive with numerals (Gregores and Suárez 1976: 144, 150, 155).

Guaraní (Gregores and Suárez 1976: 148)

(9) ipe	hikwái
duck	PL
'ducks'	

Hixkaryana and Koasati are examples of languages in which numerals do not seem to be part of the noun phrase. Although numerals in Hixkaryana are sometimes found in prenominal position (suggesting they function as noun modifiers), they typically occur ‘as sentential adverbs, related to the noun phrase, but separate from it’ (Derbyshire 1979: 44; see example below).

In Koasati the numeral does not modify the noun either but constitutes the head of a clause which (if the numeral itself is not the main predicate of the sentence) is connected to the verb of the main clause with the ‘connective suffix’ *-t* if the subject is the same (SS) and the switch reference marker *-n* if the subjects are different (DS):⁷

Koasati (Kimball 1991: 358)

(10) ná:ni-ha pokkó:l awáh tóklo-n hí:ca-li:-s
man-PL ten and two-DS see-1SG.S-PAST
‘I just saw twelve men’

The sample contains more languages with numerals that have strong verbal or nominal characteristics (e.g. Krongo, possibly Sarcee). Despite the fact that in these languages the numerals do seem to function as noun modifiers, one could debate whether or not they should be treated on a par with grammatical (i.e. non-lexical) attributive numerals. One of the most important differences is, of course, that a verbal or nominal numeral attribute is not a simple noun modifier (like ‘two’ in ‘I have two dogs’) but constitutes the head of a syntactically embedded structure which takes the referent as its argument (as in the quasi-English paraphrase: ‘I have dog[s] they/which are two’). I will show in Chapter 9 that the difference between grammatical and lexical (verbal, nominal) numerals is relevant for a description of NP-internal syntactic patterns.

It also needs to be pointed out that in Abkhaz, Chukchi, and Nivkh the numeral does not always occur as a morphologically free modifier. For example, in Chukchi numerals are bound forms, except when the NP has absolute case. Finally there are languages which properly speaking do not have attributive numerals, because the numeral forms are simply the names for gestures or (p.34) body-parts used in counting. Wambon, for instance, is a language where ‘counting is accompanied by touching the relevant body-part with the middle finger and/or index finger; the little finger of the left hand is the starting point and the nose is the turning point after which the counting goes down again via the right-hand side of the body, by prefixing *em-* “the other side” to the body part/number words’ (de Vries 1989: 41). For instance: *takhem* ‘middle finger/three’, *silutop* ‘ear/twelve’, *emsilutop* ‘ear on the other side/sixteen’. Although in Wambon these forms can also function as noun attributes (when they take the suffix *-kup* ‘added’; ‘one’, ‘two’, and ‘three’, however, have special attributive forms), it is doubtful whether this is true for all body-part count systems (cf. Greenberg 1978b: 257).

Next we take a look at some other sample languages but recall that more evidence for the classification of nominal subcategories is provided in Chapters 4 (‘Qualifying Modifiers in the Noun Phrase’) and 5 (‘Quantifying Modifiers in the Noun Phrase’).

2.2.1. Languages with singular object nouns

The marking of plural number occurs under various kinds of conditions. Here we will be mostly concerned with languages that have (a group of) nouns that are obligatory marked for plural number when modified by a free numeral higher than ‘one’ (numeral + noun + plural; no classifier).

In Abkhaz plural number is normally marked, except when the numeral is a bound form, which is attached directly to the nominal root in the case of a following non-

human noun. Notice that human and non-human nouns have different plural suffixes:

Abkhaz (Hewitt 1979: 236–7)

(11) a.	pš-y°ə(-k')	à-jyrab-c°a
	four-HUM.AX(-one)	ART-girl-PL
'(the) four girls'		

b	à-jyab-c°a	pš-y°ə(-k')
	ART-girl-PL	four-HUM.AX(-one)
'(the) four girls'		

(12) a.	pš-la-k'	b.	a-la-k° a	pš-bà
	four-dog-one		ART-dog-PL	four-BÀ
	'(the) four dogs'		'(the) four dogs'	

In Babungo and Kisi the expression of number marking is closely related to class membership; classes are distinguished on the basis of different agreement patterns.⁸ Most nouns have a double class gender in that they belong to one class in (p.35) the singular and another in the plural. Some classes remain unmarked, as in the following example from Babungo (CL = class):

Babungo (Schaub 1985: 73, 74)

(13) a.	Ø-lwáŋ	mù' b.	yí-lwáŋ	t <small>≈</small> e
	CL3-hammer	one	CL4-hammer	five
	'one hammer'		'five hammers'	

In Kisi low cardinal numerals agree in class (and number) with the head noun.

Kisi (Childs 1995: 100)

(14)	nàù-wá	á-ŋìòòŋ
	cow-SUF	PRO-two
'two cows'		

The element glossed SUF is the plural noun class suffix for inanimates and PRO is the 'prefixed noun class pronoun' (Childs 1995: 99–100).

Nouns in Bukiyp end in one of the numerous class suffixes, most of which are differentiated for singular and plural number. Numerals and other noun modifiers carry suffixes that agree in number and class with the head (Conrad 1991: 9). Compare:

Bukiyip (Conrad 1991: 59)

(15) a. bwi-yogw	beto-gw	b. bwi-yech	lowa-s
two-CL11.PL	bed-CL11.PL	two-CL13.PL	tree-CL13.PL
'two beds'		'two trees'	

In Burushaski the situation is not entirely clear. Apparently only with ‘old words’ (*les mots anciennes*) is the plural obligatory (Tiffou and Pesot 1989: 17), but there are also many nouns (e.g. for fruits and animals) which never take the plural form (Lorimer 1935–8: i. 43–5; cf. also Berger 1998: 41).

Burushaski (Lorimer 1935–8: i. 15, 187)

(16) a. hin	hir	b. altán	híri
one	man	two	men
'a man'		'the two men'	

As was already shown above Dutch nouns must appear with a plural suffix when modified by a numeral:

Dutch

(17) twee	boek-en
two	book-PL
'two books'	

In Hittite nouns with masculine and feminine gender always appear in the plural when modified by a numeral, but with neuter nouns the situation is less (p.36) straightforward: apparently they could occur with or without the plural ending (Friedrich 1974: 117; Sylvia Luraghi, personal communication).

Apart from a small group of nouns which do not appear in the plural (such as the equivalents of *mosquito*, *fish*, *duck*), Ket nouns normally carry the plural suffix when modified by a numeral (Werner 1997a: 125).⁹

Ket (personal communication)

(18) a. qok	qi'm	b. qo'm	qim-n
a	woman	ten	woman-PL
'one woman'		'ten women'	

Krongo numerals are verbal lexemes that appear in the imperfective when they modify a noun; notice also that the numeral verb agrees with the head noun in gender/number (these categories are intricately connected in Krongo; see Chapter 3):

Krongo (Reh 1985: 252)

(19) nōo-cōori nk-óotòonò

PL-house CN.PL-IMPF:be_three

'three houses'¹⁰

In Nasioi nouns denoting rational beings ('mit Vernunft begabte Wesen') are always marked for number, but nouns denoting non-rational and inanimate entities ('vernunft- oder leblose Dingen') do not show any sign of number. Note, however, that nouns in definite NPs always carry a suffix simultaneously marking definiteness and number (in definite NPs numerals agree in class with the head noun):

Nasioi (Rausch 1912: 108)

(20) bau bénamo

taro three (= absolute/indefinite form of 'three')

'three taros'

(21) bau-ni békupi

taro-DEF.PL three (= class 6 form of 'three', i.e. for fruits)

'the three taros'

Thus the transnumeral character of the non-rational noun only shows in the indefinite NPs (cf. Rausch 1912: 115). As a matter of fact the sample contains several languages (e.g. Chukchi, Koasati) in which number is expressed in a portmanteau element which simultaneously expresses such values as gender or noun class, definiteness and/or case. In other words, there are languages in which (p.37) the expression of one notion (such as definiteness or case) necessarily involves the expression of another notion (number).¹¹

In Ngiti almost all nouns are unmarked for number; only nouns denoting humans as well as compound forms whose second part is (originally) a 'human' noun are marked for number ((i)nzo is the suppletive plural of *ngba* 'child'):

Ngiti (Kutsch Lojenga 1994: 340, 355)

(22) ifɔ ikyì-akpá-nzo

four cow-male:PL-children

'four young bull-calves'

(23) àrù tsu

eight tree

'eight trees'

Nouns in West Greenlandic are also marked for number when modified by a numeral:

West Greenlandic (Fortescue 1984: 135)

(24) puisi-t quilit

seal-PL ten

'ten seals'

Although the evidence is inconclusive, it seems that nouns in Pipil take a plural affix when modified by a numeral. The plural suffix used in this example is the reduplicated form of *-tsin* 'diminutive' or 'reverential' (Campbell 1985: 51–2).

Pipil (Campbell 1985: 104)

(25) ne ye:y pipu-tsi-tsín

the three boy-PL-DIM

'the three little boys'

In Sarcee only human nouns and the nominals *tłí*(*tc’á*;) 'dog' and *ıstħi* 'horse' (which derives from 'dog') may occur with an element that seems to combine with expressions of plurality or collectivity.

Sarcee (Cook 1984: 65)

(26) gūnìsnóní-ká tłí-ká

nine-PL dog-PL

'nine dogs'

As the example shows, the *-ká* suffix may additionally appear on the numeral 'as a kind of agreement' (Cook 1984: 65).¹²

(p.38) In Tamil only nouns of the 'rational' class are always marked for number (except *peeru* 'person'), other nouns usually remain unmarked:

Tamil (Asher 1982: 61)

(27) rantu pustakam

two book:NOM.SG

'two books'

2.2.2. Languages with set nouns

This section will only be concerned with languages in which nouns are normally not marked for number while being in a direct construction with a free numeral modifier (numeral + noun; no plural, no classifier). A small or large group of such nouns are also attested in some of the languages mentioned in section 2.2.1. (e.g. Hittite and Ngiti),

but the list is much longer; in fact, from a cross-linguistic perspective number marking (with or without an attributive numeral) seems to be the exception rather than the rule.

In the following languages nouns are unmarked for number when modified by a free numeral: Basque, Berbice Dutch Creole (see 4.2.1.2), Chukchi (see below), Galela, Georgian, Guarani, Gude,¹³ Hittite, Hixkaryana (but see 2.2), Hungarian, Ika, Kayardild, Lango,¹⁴ Hmong Njua (see 4.2.1.2),¹⁵ Nahali, Ngalakan (no examples, but see Merlan 1983: 39, 53, 56, 90), Nunggubuyu (no example, but see note 7), Oromo, Sumerian (Thomsen 1984: 59 f.), Wambon, and Tsou.

Nouns in Nivkh are not marked for number in the context of quantification, but when numerals do not combine (or rather amalgamate) with a classifier (see 2.2.3), they precede the head noun as part of a polysynthetic construction according to Mattissen and Drossard (1998: 50). Chukchi is also a special case, because numerals can only appear as free modifiers when the noun (phrase) is in the absolute case (i.e. subject of intransitive or object of transitive verb); otherwise they are incorporated.

Chukchi (Masja Koptjevskaja-Tamm, personal communication):

(28)	mytlynjen	jar-at
	five	house-PL.ABS
'five houses'		

(p.39) I have nevertheless assumed that Chukchi has set nouns because plural number (if it is that; see Chapter 3) is ‘neutralized’ in all cases other than the absolute (Bogoras 1922: 694; Comrie 1981a: 246). Recall that in Abkhaz, too, the noun is unmarked for number in the case of a bound numeral (see example above). Below are examples from Basque, Berbice Dutch Creole, Galela, Georgian, Guarani, Gude, Hixkaryana, Hungarian, Ika, Kayardild, Lango, Nahali, Nivkh, Oromo, Sumerian, Tsou, and Wambon.

(29)	hiru	neska
	three	girl
'three girls' ¹⁶		

Berbice Dutch Creole (Kouwenberg 1991: 168)

(30)	twe	jermatoko
	two	woman = child
'two daughters'		

Galela (van Baarda 1891: 13)

(31)	o	tahoe	motohha
	ART	house	five
'five houses' ¹⁷			

Georgian (Fähnrich 1986: 158)

(32) ori mçerali

two writer

'two writers'

(p.40) Guaraní (Gregores and Suárez 1967: 150)

(33) ko mokõi lo mitã

this two the boy

'these two boys'

Gude (Hoskison 1983: 51)

(34) mbusæ pu'

pumpkin ten

'ten pumpkins'

Hixkaryana (Derbyshire 1979: 44)

(35) asak kanawa

two canoe

'two canoes'

Hungarian (Moravcsik 1994: 8)

(36) két lány

two girl

'two girls'

Ika (Frank 1990: 32)

(37) mouga tšeirua

two man

'two men'

Kayardild (Evans 1995: 235)

(38) kiyarrng-ka yarbud-a ngarnal

two-NOM	meat-NOM	white_cockatoo: NOM
(Number)	(Generic)	(Specific)
‘two white cockatoos’		

Lango (Noonan 1992: 167)

(39) gúlú áryô
pot two
‘two pots’

Nahali (Kuiper 1962: 265)

(40) tîr tâñâ
two son
‘two sons’

Nivkh (Gruzdeva 1998: 24)¹⁸

(41) nax pitxan̩-ø
six book-NOM
‘six books’

(p.41) Oromo (Stroomer 1987: 59)

(42) gaala lamaani
camel two
‘two camels’

Sumerian (Thomsen 1984: 83)

(43) abgal imin-e
sage seven-ERG
‘the seven sages’

Tsou (Szakos 1994: 163)

(44) euso 'o kueai
two ART car
‘two cars’ ¹⁹

Wambon (de Vries 1989: 42)

(45) ap	ilumtakhemo
house	three
‘three houses’	

2.2.3. Languages with sort nouns

In the following languages the numeral combines with an extra element, the so-called sortal classifier, and the noun is not marked for number (numeral + noun + classifier; no plural):²⁰

Mandarin Chinese (Li and Thompson 1989: 105)

(46) nèi-liù-běn	shū
that-six-CLF	book
‘those six books’	

Burmese (Okell 1969: 209)

(47) hkwei	hnă	kauñ
dog	two	CLF
‘two dogs’		

Korean (Lee 1989: 55)

(48) mat	du	mali
horse	two	CLF
‘two horses’		

(p.42) Nivkh (Comrie 1981a: 269)²¹

(49) qan	mor
dog	two:
CLF(animal)	
‘two doss’	

Nung (Saul and Freiberger Wilson 1980: 23)

(50) slóng	tú	lúhc
------------	----	------

two CLF child

'two children'

Vietnamese (Nguyễn Đình Hoà 1987: 785)

(51) ba quyển sách

three CLF book

'three books'

Although most of these languages are also deemed to have some rare, optional kind of plural marking, I will argue in section 5.2.1.3 that these so-called plural number markers are better categorized as quantifiers (and in some cases as *nominal aspect markers*—more specifically *collective aspect markers*—which are typically associated with *set nouns*; cf. section 4.2.1.2).²²

2.2.4. Languages with flexible nouns

Let us finally turn to the languages without a distinct class of nouns: Samoan (type 1: V/N/A), Quechua, Turkish, and Hurrian (type 2: N/A).²³ As may be expected, flexible lexemes of either type (V/N/A and N/A) are not specified for such noun specific categories as number and gender or noun class, i.e. flexible nouns are transnumeral (Chapter 5) and are not divided into different genders (Chapter 3). The transnumeral character of Samoan lexemes (V/N/A) is illustrated in the following example (see also 5.3.1 on the expression of cardinality in Samoan):

Samoan (Mosel and Hovdhausen 1992: 115)

(52) E tolu tusi na maua

GENR three letter PAST get

'[She/he] got three letters'

[lit. 'The letters she/he got are three']

(p.43) In Imbabura Quechua the head of the NP may optionally appear with the plural suffix *-kuna* (which also marks plural number on pronouns) when modified by a numeral:²⁴

Imbabura Quechua (Cole 1982: 128)

(53) ishkay wasi(-kuna)-ta chari-ni

two house(-PL)-ACC have-1

'I have two houses'

Speiser (1941: 101) regards the suffix *-na* as a plural marker in Hurrian, but he notes that it is often omitted, especially 'if the desired concept of plurality is marked elsewhere in the sentence' (Speiser 1941: 101). The same morpheme has, however, also

been called an anaphoric element ('ones', Speiser 1941: 102), an article, and a 'relator'. Apparently *-na*, which (like Quechua *-kuna*) occurs between the lexeme and the case ending, is incompatible with possessive suffixes and does not occur with instrumental *-ae* and essive *-a* (Gernot 1995: 114).

In Turkish, too, nouns are both 'devoid of grammatical gender' and trans-numeral or 'numerically neutral' (Lewis 1967: 25–6). The so-called plural suffix *-ler/-lar* is not obligatory and is normally absent when the noun is modified by a numeral (see also Underhill 1976: 125).²⁵

Turkish (van Schaik 1996: 95)

(54) bu	iki	genç	cocuk
	this	two	young child
'these two young children'			

Historically *-ler/-lar* is classified as a collective suffix, which in Old Turkish (sixth/seventh century) only occurred with words referring to sovereigns and older female relatives. Later, it was first employed to mark the (collective) plurality of nouns denoting animate entities, then of nouns in general (Grönbech 1936: 61 f.; Schroeder 1999: 112). Interestingly, as in the case of Quechua, there is a co-occurrence restriction in that in a possessive construction the so-called plural suffix can only occur once, even when both the possessor and the possessed are plural. Consequently, a noun with a third person plural possessive suffix such as *oda-lar-I* [room-PL-3:POS.SX] can mean 'his rooms', 'their room', or 'their (p.44) rooms' (Underhill 1976: 92–3; see also Lewis 1967: 39).²⁶ Note also that both in Quechua and in Turkish the so-called nominal plural marker serves a double purpose in that it also marks plural number in the main predicate.²⁷

2.3. Towards a comprehensive classification of nominal subcategories

Before I proceed, it may be useful to reiterate that referents of NPs are not objects in the real world, but rather mental constructs that are created, stored, and retrieved in the minds of the speech participants. It is important to keep this in mind, since this allows for possible discrepancies between (linguistic) properties of discourse referents and (ontological) properties of their real-world counterparts or *Sein*-correlates (if they exist). It will appear below that this holds especially true with respect to the features Shape and Homogeneity.

2.3.1. Nominal subcategories (reprise)

In this section I will elaborate on the morpho-syntactic properties of the three noun types distinguished above, which all have to do with quantification. Furthermore I will attempt to establish a relationship between these morpho-syntactic properties and semantic characteristics of the various noun types and argue that there is actually a fourth nominal subcategory that is used to refer to discrete spatial objects (besides the three subcategories introduced in section 2.2: *singular object nouns*, *set nouns*,

and *sort nouns*). For the sake of the argument I will restrict myself to nouns in four genetically unrelated languages that were chosen for illustrative purposes: Dutch, Oromo, Thai, and Yucatec Maya. The last two languages are not in the sample. Thai was selected because the article by(p.45) Hundius and Kölver (1983) is one of the very few studies that explicitly deal with noun semantics in a classifier language and Yucatec Maya is included in the discussion below because it uses a nominal subcategory that is not found in any of the other languages in the sample.

2.3.1.1. Dutch versus Oromo, Thai, and Yucatec Maya

It has already been established that Dutch nouns can be in a direct construction with a numeral and must appear in the plural form when modified by a numeral ($n > 1$). Comparing Dutch with Oromo, Thai, and Yucatec Maya we find that only in Dutch an NP headed by a *bare* noun must refer to a single individual. When reference is made to more than one individual object, the noun must be suffixed with a plural marker. In other words, plural marking is obligatory, both with and without a numeral modifier (where $n > 1$):

Dutch

(55) a. de auto	b. de auto-s
the car	the car-PL
'the car'	'the cars'

Dutch

(56) a. *de twee auto	b. de twee auto-s
the two car	the two car-PL
	'the two cars'

I call such nouns (which are typically found in the European languages) *singular object nouns* (section 2.2). By contrast bare nouns in Oromo, Thai, and Yucatec Maya are *transnumeral*, i.e. these nouns are neutral with respect to the number of individuals they denote (cf. Greenberg 1972: 24). Consider the following citations and examples:²⁸

Oromo (Andrzejewski 1960: 71)

The vast majority of [nouns] are associated with NEITHER PLURALITY NOR SINGULARITY, i.e. the forms themselves give us no information as to whether what is denoted by them is one or more than one. When such forms are used, only the context can provide us with information about the number of what is denoted.

Oromo (Stroomer 1987: 74)

(57) a. nama	'man/men'
b. śimp 'irree	'bird/birds'

(p.46) *Thai* (Hundius and Kölver 1983: 181–2)

Thai nouns do not in themselves contain any numerical or referential indications ... they are purely *conceptual labels* which, in order to be appropriately related to objects of the non-linguistic world, always and in principle stand in need of interpretation which has to be inferred from both linguistic and non-linguistic context.

Thai (Hundius and Kölver 1983: 172, 181)

(58) a.	rōm	sǐi-khǐaw				
	umbrella	green				
		'green umbrella(s)'				
b.	plaa	wâaj	júu	naj	mêe	náam
	fish	swim	remain	in	river	
		'fish swim in rivers' [generic]				
		'the fish/fishes swims/swim in a river/rivers'				
		'a fish/fishes swim(s)/swam in the river/a river/...etc.'				

Yucatec Maya (Lucy 1992: 46–7, 55)

Yucatec noun phrase heads have truly *neutral* form with respect to number the lexical nouns can be used just as well with a multiplicity of referents as with one referent.

In general,... grammatical number is not syntactically deductible in Yucatec by reference to underlying lexical content.

2.3.1.2. Dutch and Oromo versus Thai and Yucatec Maya

Dutch and Oromo differ from Thai and Yucatec Maya in that in the first two languages the noun can be in a direct construction with a numeral.

Dutch

(59)	twee	boek-en				
	two	boek-PL				
		'two books'				

Oromo (Stroomer 1987: 59)

(60)	gaala	lamaani				
	camel	two				
		'two camels'				

It has been noted before that in Oromo but not in Dutch the noun remains unmarked for number when it is modified by a numeral.²⁹ I call nouns of the Oromo type *set nouns*, because they seem to denote a set of individuals. A set can have any cardinality: it may contain just one individual (singleton set) or it may (p.47) consist of more individuals (multiple set). I will return to set nouns in Chapter 4. Here are some more examples:

Oromo (Stroomer 1987: 75, 76, 84)

- | | | |
|---------|----------|----------------|
| (61) a. | farda | 'horse/horses' |
| b. | saree | 'dog/dogs' |
| c. | nad'eeni | 'woman/women' |

In Thai and Yucatec Maya, on the other hand, nouns cannot be in a direct construction with a numeral; in these languages the numeral must combine with a classifier.

Thai (Hundius and Kölver 1983: 166):

[Thai nouns] purely denote *concepts* and, for this reason, are incompatible with direct quantification.

Thai (Hundius and Kölver 1983: 172)

- | | | |
|------|-------------------|--------------------------------|
| (62) | rôm | săarn khan |
| | umbrella | three CLF:long, handled object |
| | 'three umbrellas' | |

Yucatec Maya (Lucy 1992: 43)

Enumeration with Yucatec numerals requires the insertion of a numeral classifier between numeral and lexical noun (if there is one) which provides information about the unit being counted.

Compare:

Yucatec Maya (Lucy 1992: 74, 2000: 329):³⁰ a/one-CLF banana

- | | | | |
|---------|------------|-------|---|
| (63) a. | 'un-tz'iit | há'as | 'one/a 1-dimensional banana (i.e. the fruit)' |
| b. | 'un-wáal | há'as | 'one/a 2-dimensional banana (i.e. the leaf)' |
| c. | 'un-kúul | há'as | 'one/a planted banana (i.e. the plant/tree)' |
| d. | 'un-kúuch | há'as | 'one/a load banana (i.e. the bunch)' |
| e. | 'um-p'iit | há'as | 'one bit banana (i.e. a bit of the fruit)' |

2.3.1.3. Thai versus Yucatec Maya

Although both in Thai and in Yucatec Maya the noun cannot be in a direct construction with a numeral, it would be wrong to say that nouns in these languages belong to the

same kind of nominal subcategory. This can be demonstrated by the fact that they employ different kinds of classifiers. Traditionally only two kinds of classifiers are distinguished: *mensural* and *sortal* (or *numeral*) ([p.48](#)) *classifiers*. *Mensural classifiers* indicate size, volume, or weight and typically occur in combination with nouns denoting non-discrete spatial entities (masses): ‘a liter of wine’, ‘a bag of flour’, ‘a pound of cheese’, ‘a cup of tea’. *Sortal classifiers*, on the other hand, are typically used in connection with discrete objects. They do not indicate the volume or size, but may involve many different kinds of notions (notably shape). Compare:³¹

Sortal classifier:

Thai (Gandour et al. 1984: 466, 455)

-
- (64) a. thian sii lêm
candle two CLF: long, pointed object
'two candles'
-
- b. pèt hâa tua
duck five CLF:body
'five ducks'

Mensural classifier:

Thai (Hundius and Kölver 1983: 168, 170)

-
- (65) a. dinniaw sääm kõon
clay three lump
'three lumps of clay'
-
- b. náamtaan sääm thúaj
sugar three cup
'three cups of sugar'

I call nouns that occur with sortal/numeral classifiers *sort nouns*; and nouns that occur with mensural classifiers (“measures”) *mass nouns*.³² Sortal classifiers can be further divided into classifiers that are used to count single entities (sometimes called *common classifiers*) and classifiers that are used for counting discrete entities in groups, the group or *collective classifiers*. In many languages the collective classifier is the diachronic source of nominal (i.e. collective) aspect and plural markers (section 4.2.1.2). The following examples of collective classifiers are from Burmese:

Burmese (Okell 1969: 211)

-
- (66) pyà hnă ouñ
bee two swarm
'two swarms of bees'

(p.49) Burmese (Okell 1969: 211)

(67)	pāñ	hnă	sì
	flower	two	bunch
'two bunches of flowers'			

Besides sortal and mensural classifiers there is another major kind of classifier type which is attested in Yucatec Maya (Lucy 1992: 74):

Outside of the restriction on compatibility with other classifiers, little in the grammar of Yucatec appears to hinge on, or correlate with, this ‘sortal’ ... versus ‘mensural’ distinction, and it is difficult to know what status it should be given.

This suggests that in Yucatec Maya there is no clear distinction between *sort nouns* and *mass nouns*. I will call such nouns *general nouns*, and the classifiers that are used with these nouns *general classifiers*.

The difference between Thai and Yucatec Maya nouns is also illustrated by the following examples from Thai, which should be compared with the Yucatec examples in (63):

Thai (Hundius and Kölver 1983: 186–7)

(68) a.	mámûaŋ	sääm	lūuk
	mango	three	CLF:fruit
'three mangoes'			
b.	mámûaŋ	sääm	baj
	mango	three	CLF:leaf
'three mangoes'			

The phrases in (68a) and (b) are stated to be fully equivalent. So as to be able to express the meaning *three mango leaves* in Thai, [t]here is no other way than first to form the corresponding nominal compound *bajmámûaŋ* “mango-leaf”, which then, in its turn, is appropriately classified under *baj* (ibid.):

Thai (Hundius and Kölver 1983: 187)

(69)	baj +	mámûaŋ	sääm	baj
	leaf +	mango	three	CLF:leaf
'three mango leaves'				

The examples in (63) show that Yucatec nouns are even less specified in terms of lexically coded information than Thai nouns. Whereas speakers of Thai must form a compound (*baj + mámûaŋ* = leaf + mango), speakers of Yucatec Maya simply use the

same noun (*há*’ as ‘banana’) in combination with a different classifier (namely *-wáal*): ‘*un-wáal há*’ as one-Clf banana ‘one/a two-dimensional banana (i.e. the leaf)’.

(p.50) 2.4. A new classification of nominal subcategories

In the previous sections I have argued on the basis of evidence from four unrelated languages (Dutch, Oromo, Thai, Yucatec Maya) that at least four different kinds of nouns are used to refer to discrete spatial entities in the external world (see note 1):

- *singular object nouns* (Dutch): the noun is in a direct construction with the numeral; plural marking is obligatory, both with and without a numeral;
- *set nouns* (Oromo): the noun is in a direct construction with the numeral; so-called number marking (if available at all; see Chapter 4) is absent when the noun is modified by a numeral;³³
- *sort nouns* (Thai): the noun is not in a direct construction with the numeral; instead the numeral combines with a *sortal classifier* and number marking is absent, both with and without a numeral;
- *general nouns* (Yucatec Maya): the noun is not in a direct construction with the numeral; instead the numeral combines with a *general classifier* and number marking is absent, both with and without a numeral (but see also note 30).

In the next section I will argue for a more comprehensive, semantically orientated classification of nominal subcategories on the basis of two features: Shape and Homogeneity.

2.4.1. The feature ‘Shape’

One of the most perspicuous properties of nouns in Thai and Yucatec Maya is that they can only be numerated when the numeral combines with a classifier. Apparently this has to do with properties of the noun rather than those of the numeral; according to the sources (Hundius and Kölver 1983 on Thai, Lucy 1992 on Yucatec Maya) the meaning definitions of nouns in these languages do not include the notion of spatial boundedness or discreteness. Since only discrete entities can be numerated, one may assume that this is the reason why in languages such as Thai and Yucatec Maya the numeral must combine with a *classifier*, which functions as a kind of *individualizer* (cf. Lyons 1977: 462).³⁴

Thai (Hundius and Kölver 1983: 166; the quotation is repeated for convenience):

[Thai nouns] purely denote concepts and, for this reason, are incompatible with direct quantification.

(p.51)

Table 2.2. – Shape and + Shape nominal subcategories

- | | |
|---------|-------------------------------------|
| – Shape | sort nouns and general nouns |
| + Shape | singular object nouns and set nouns |

Yucatec Maya (Lucy 1992: 83, 43)

Interpretatively, in Yucatec all nouns ... are neutral with respect to logical unit or shape.³⁵

In other words, in Thai and Yucatec Maya nouns that are used to refer to discrete spatial objects designate properties that are not characterized as having a definite outline (Shape) in the spatial dimension. Therefore I propose to use the feature ‘Shape’ to distinguish between *singular object nouns* and *sort nouns* on the one hand (both + Shape), and sort nouns and general nouns on the other (both –Shape); see Table 2.2. Thus, as a rule the value –Shape correlates with the use of classifiers (or: individualizers) and, conversely, the value +Shape correlates with the absence of classifiers (but see section 4.3.5 on Hmong Njua).

2.4.2. The feature ‘Homogeneity’: masses and collectives

It appears that there is also a semantic feature in the meaning definition of nouns which enables us to distinguish between nominal subcategories on the vertical axis in Table 2.2, i.e. between (i) *singular object nouns* and *sort nouns* on the one hand, and between (ii) *set nouns* and *general nouns* on the other, namely ‘Homogeneity’ (also ‘likepartedness’ or ‘dissectiveness’; Goodman 1966). To make clear what ‘Homogeneity’ means in this context, I shall briefly discuss two other major nominal subcategories, namely *mass nouns* and *collective nouns*. As these nouns are not used to refer to a single individual, they were largely ignored in the previous sections.

(p.52) Mass nouns such as English ‘water’ and ‘gold’ define [+Homogeneous] entities because they are cumulative (or agglomerative) and—up to a point—dissective (Goodman 1966). If we add some water to a liter of water we still refer to it as water (cumulative/agglomerative); after we have drunk some of the water that is contained in a glass, the glass will still contain water (dissective). Singular object nouns like Dutch *fiets* ‘bicycle’ and sort nouns like Thai *rôm* ‘umbrella, umbrellas’, on the other hand, define non-homogeneous entities: we cannot refer to something as ‘a bicycle’ or ‘an umbrella’ if it is something more or less than a (one) bicycle or umbrella (as when e.g. essential parts are missing; at best we could speak of a broken or defective bicycle or umbrella). We could say that a referent defined by a mass noun consists of portions, whereas an entity defined by a singular object noun or a sort noun consists of parts or components. Adding portions gives you more of the same *mass entity*, but adding parts does not give you more *sort entities*.

To sum up, using the features Shape and Homogeneity to characterize *mass nouns*, *singular object nouns*, and *sort nouns* gives the following distribution of features:

Mass nouns:	–Shape, +Homogeneity
Singular object nouns:	+Shape, –Homogeneity
Sort nouns:	–Shape, –Homogeneity

The feature Homogeneity also distinguishes *mass nouns* from *general nouns* and *set nouns*. Members that belong to the latter two categories are *neutral* with respect to the feature Homogeneity. Regarding *general nouns*, recall that Yucatec Maya does not

seem to make a distinction between *sortal classifiers* (which occur with *sort nouns*) and *mensural classifiers* (which combine with *mass nouns*). From this it was inferred that there is no distinction between *sort nouns* (–Shape, –Homogeneity) and *mass nouns* (–Shape, +Homogeneity) in Yucatec Maya, which suggests that the feature Homogeneity is irrelevant for *general nouns* (–Shape).

Set nouns (+Shape), which designate a spatially bounded property, are also *neutral* with respect to the feature ‘Homogeneity’, in that a set entity may or may not be homogeneous. If the set comprises more than one individual, it can ultimately be divided into as many singleton sets as there are individuals. In the case of a singleton set, however, the space for which the property holds cannot be further divided. Thus the feature ‘Homogeneity’ is basically irrelevant for *general nouns* and *set nouns*:

At this point five nominal subcategories have been characterized in terms of Shape or ± Shape, ± Homogeneity:

General nouns:	–Shape
Sort nouns:	–Shape, –Homogeneity
Mass nouns:	–Shape, +Homogeneity

(p.53)

Set nouns:	+Shape
Singular object nouns:	+Shape, –Homogeneity

Finally, *collective nouns* (+Shape, since they can be in a direct construction with a numeral: e.g. *two families*) differ from *mass nouns* (–Shape) in that the former but not the latter are obligatorily marked for plural number (unless creating a shift in meaning, as in *two wines* i.e. ‘two kinds of wine’): *a/one family, two families* vs. * *a/one gold, two golds*. As has been mentioned earlier, *mass nouns* also require the occurrence of a mensural classifier when they are quantified: e.g. *two bars of gold, five liters of water*. But *collective nouns* and *mass nouns* are similar in that the space for which the property holds is divisible (+Homogeneity).

Collective nouns designate a property of several discrete entities that are conceived as a unit. But when, for instance, someone in a family dies, the others are still ‘family’, and when we take one flower out of a bunch, the remaining flowers are still members of (part of) a *bunch*. Since there is a lower limit on the number of individuals that make up a collective entity, it is perhaps better to say that a collective (or a mass) is *agglomerative* in that it can be expanded with members (or portions) of the same kind, without affecting the number of the NP (although this does change the size of the *collective*, or the size, weight, or volume of the *mass*). For instance, if a family is expanded (as when e.g. a child is born into the family), this does not change the number of the collective ‘family’, but only the number of individuals that make up that family.

2.4.3. Defining the nominal subcategories

Six nominal subcategories can now be defined on the basis of the features Shape and

Homogeneity (see Fig. 2.1).³⁶ If the property designated by a noun is coded as having shape (+Shape), this means that the property (and, by extension, the referent of the NP) is characterized as having a definite outline in the spatial dimension; hence set nouns, singular object nouns, and collective nouns can all be in a direct construction with a cardinal numeral. If the property designated by a noun is coded as being homogeneous (+Homogeneity), this means that the space for which this property holds (and, by extension, the referent of the NP) is characterized as being agglomerative. In other words, the referent of an NP headed by a noun that is coded as being homogeneous consists of *portions* (of a mass) or *members* (of a collective). *General nouns* and *set nouns* are neutral with respect to the feature Homogeneity. It is now possible to characterize six (p.54)

Space	–Homogeneity	+Homogeneity
–Shape	general noun	
	sort noun	mass noun
+Shape	set noun	
	singular object noun	collective noun

Fig. 2.1. Cross-linguistic classification of major nominal subcategories
nominal subcategories as follows:³⁷

- *General noun* (–Shape): the noun designates a property that is only characterized as *not* having a definite outline;
- *Sort noun* (–Shape, –Homogeneity): the noun designates a property that is characterized as *not* having a definite outline and as being *non-agglomerative*;
- *Mass noun* (–Shape, +Homogeneity): the noun designates a property that is characterized as *not* having a definite outline and as being *agglomerative*;
- *Set noun* (+Shape): the noun designates a property that is only characterized as having a definite outline;
- *Singular object noun* (+Shape, –Homogeneity): the noun designates a property that is characterized as having a definite outline and as being *non-agglomerative*;
- *Collective noun* (+Shape, +Homogeneity): the noun designates a property that is characterized as having a definite outline and as being *agglomerative*.

It is not clear why the feature Shape is relevant for all noun types, whereas the feature Homogeneity helps to define only four nominal subcategories. Yet this classification confirms Friedrich's observation that 'the category of shape appears to be a typological universal in grammar ..., and of not inconsiderable significance for a theory of semantics in grammar' (Friedrich 1970: 380). The fact that in this classification the feature 'Shape' is more important than the feature 'Homogeneity' may also have to do with the idea that spatial orientation is (p.55) primary in human cognition and links

up nicely with observations in other grammatical domains (more on this in Chapter 7).³⁸

Recall also, once again, that this is a linguistic rather than an ontological classification in that it is only based on morpho-syntactic and semantic properties of nouns.³⁹ As for the terminology, I could also have used labels such as type A and type B etc., but for mnemonic reasons I have tried to use names which say something about the semantics of each nominal subcategory. Note furthermore that, since this is a linguistic and not an ontological classification, there is in principle no direct relationship between noun type and (real world) entity type. This is, of course, precisely the reason why it is possible for different noun types to be used across the world's languages to refer to entities whose *Sein*-correlate in the external world is a single and discrete spatial object: singular object nouns, set nouns, sort nouns, and general nouns. In other words, speakers of a different language do not necessarily use a member of the same nominal subcategory when they talk about, for instance, a knife.⁴⁰ As I have pointed out before, the fact that this is possible has to do with the status of referents. In section 1.5.5 it was established that the immediate referent of a noun phrase is a mental rather than a physical entity, because it is also possible to refer to persons, animals, plants, objects that do not exist: as soon as an imaginary entity has been introduced into the discourse world ('I read a book about *the Holy Grail/the unicorn/the yeti/Captain Spock/...*'), it exists as a discourse referent (Karttunen 1976) and can be referred to again just like a referent that does have a *Sein*-correlate in (p.56) the external world.⁴¹ It seems, then, that from a linguistic point of view there is no difference between referents with and referents without a *Sein*-correlate in the external world (cf. Rapaport 1985/6). All referents have a purely mental status, but only a subset of all referents also have a physical counterpart in the external world.

2.5. Nominal subcategories: beyond the noun phrase

In the previous section I have used certain morpho-syntactic features as a heuristic device to propose a semantic classification of the major nominal subcategories (Fig. 2.1). On the assumption that the classification of nouns in terms of the two features 'Shape' and 'Homogeneity' truly captures the main semantic differences between nouns as they occur in the languages of the world, this feature-based approach to nominal subcategories also makes it possible to account for such widely different phenomena as the meaning of incorporated and predicate nouns (2.5.1) and certain instances of so-called number discord (for reasons that will become clear later, the latter issue is discussed in Chapter 4).

2.5.1. The semantics of predicate and incorporated nouns

It has often been observed that when nouns are incorporated to become part of a derived verbal predicate, this involves some kind of semantic change in that the incorporated noun is said, for instance, to be non-individuated or to have lost its individual salience both semantically and syntactically.⁴² This (admittedly vague) semantic difference between incorporated and non-incorporated nouns may be understood better if we accept that verbal predicates cannot designate properties (or relations) that pertain to both the temporal and the spatial dimension. If this premiss is accepted, then it must follow that an incorporated noun (e.g. 'baby' in *[to] babysit*) does not designate a property in the spatial dimension, hence the property it designates

cannot be characterized in terms of the spatial features Shape and Homogeneity. Obviously, if an incorporated noun does not designate a property in the spatial dimension, it is impossible to indicate how such a property is represented in that dimension. Thus, the semantic difference between an incorporated and a non-incorporated noun is due to the fact that incorporated nouns are maximally underspecified nominal predicates, which in turn, is a consequence of the fact that incorporated nouns do not designate properties in the spatial dimension.

(p.57) The same probably holds for predicate nouns, which, at least in Dutch, are said ‘not to individuate’ (Geerts et al. 1984: 145). Compare (note that *soldaat-en* is spelled *soldaten*):

Dutch

- | | |
|---|--|
| (70) a. Jan is een soldaat | |
| Jan is a soldier | |
| ‘Jan is a soldier’ | |
| <hr/> | |
| b. Jan en Piet zijn twee goede soldaat-en | |
| Jan and Piet are two good soldier-PL | |
| ‘Jan and Piet are two good soldiers’ | |

Dutch

- | | |
|--|--|
| (71) a. Jan is soldaat | |
| Jan is soldier | |
| ‘Jan is a soldier’ | |
| <hr/> | |
| b. Jan en Piet zijn soldaat | |
| Jan and Piet are soldier | |
| ‘Jan and Piet are soldiers’ | |
| <hr/> | |
| c. *Jan en Piet zijn een goede soldaat | |
| Jan and Piet are a good soldier | |

In (70a-b) *soldaat* ‘soldier’ heads an NP, as is indicated by the indefinite article *een*, and by the occurrence of such modifiers as *twee* (a numeral) and *goede* (an adjective). In (71a-b) the bare noun *soldaat* is the sentence predicate and as such it seems to have lost a number of nominal properties: it is not inflected for number (whatever the number of the subject NP), and (71c) illustrates that it cannot combine with modifiers either (here, article and adjective).

2.6. Nominal subcategories as different Seinsarten

We have seen above that nouns can be distinguished in terms of the spatial features Shape and/or Homogeneity into six major noun types. Verbs, by contrast, can in

principle all be characterized in terms of the two typically temporal features Beginning and Ending. With these features we can define what are commonly called aspectual distinctions:

(72) ----- = imperfective	(-Beginning/-Ending)
Ti > ----- = ingressive	(+Beginning/-Ending)
----- < Tj = egressive	(-Beginning/+Ending)
Ti >----- < Tj = perfective	(+Beginning/+Ending) ⁴³

(p.58) Let us take as an example the verb ‘to sleep’. Since these distinctions are hardly coded on verbs in English, I will use paraphrases to explain the meaning. If the speaker uses the verb in the perfective form he emphasizes the temporal boundedness of the sleeping event. With sleep in the egressive form he stresses the ending (‘to stop sleeping’ =to wake up), whereas sleep + ingressive aspect underlines the beginning of the sleeping event (i.e. ‘to fall asleep’). Sleeping in the imperfective form, finally, does not draw attention to the beginning or the ending but to the occurrence of the event as such. In many languages imperfective aspect is used to provide a background for a more central event, as in e.g. ‘While she was sleeping, her children set fire to the curtains.’ In other words, the same property (‘sleep’) can be represented in at least four different ways in terms of the features Beginning and Ending. Especially the choice between perfective and imperfective is often a matter of pragmatics in that it is determined by what the speaker wishes to emphasize. For instance, the English sentence *I stood there for an hour* can be translated in Russian as *ja stojal tam cas* (with the verb in the imperfective form) or as *ja postojal tam cas*, i.e. with the verb in the perfective form (from Comrie 1976: 4, 16–17). According to Comrie the last sentence (with *postojal*) implies that the waiting was not experienced as lasting long whereas the first sentence (with *stojal*) is neutral in this respect. In other words, the same event in the physical world can be represented in different ways. Of course, this also shows that sentences (or rather predication), just like NPs, do not refer to entities (here: events) in the real world, but rather to mental constructs of events which may or may not have a *Sein*-correlate, or rather *Aktion*-correlate, in the external world.

The four aspectual distinctions can also be represented as in Table 2.3. I have deliberately avoided the phrase ‘verbal aspects’ in Table 2.3, because depending on the way the distinctions are coded they belong to different fields in verb semantics. When they are expressed by inflectional morphology they are usually called verbal aspects, but when these aspectual distinctions are part of the lexical meaning of a verb, i.e. when they are morphologically invisible, they are usually studied under the heading of *Aktionsart* (for want of a good English term; it could be translated as ‘mode of action’ or ‘type of State-of-Affairs’ (cf. Comrie 1976: 6–7; Dik 1997: 105–26). To give an example, the perfective (more precisely, momentaneous) character of a verb like ‘to hit’ (*the arrow hit the target*) and the

Table 2.3. Verbal categories

Time	-Beginning	+Beginning
-Ending	imperfective	ingressive
+Ending	egressive	perfective

(p.59) telic character of a verb such as ‘to cross (the street)’ or ‘to reach (the top)’ belong to the study of *Aktionsart* and not verbal aspect.⁴⁴

The reason for briefly mentioning the difference between covert (lexical) and overt (inflectional) marking of verbal categories is that we can do the same in the area of noun semantics. To the extent that the features Shape and Homogeneity are part of the lexical meaning of the noun they could be studied in the context of *Seinsart* (i.e. mode of being), to stick to the German terminology in this area of lexical semantics. And when these features are overtly coded by inflectional morphology we could speak of nominal aspect (nominal aspect marking will be dealt with in section 4.2). Thus, *Seinsart* deals with the covert (lexical) coding of the way a nominal property is represented in the spatial dimension in terms of the features Shape and/or Homogeneity and *nominal aspect* is reserved for only the overt inflectional expression of Shape and/or Homogeneity.

It is also important to emphasize at this point that languages do not so much differ in the *kind* of nominal properties they predicate of entities, but rather in the way the meaning definition of the noun specifies how the property is represented in the spatial dimension in terms of the features Shape and Homogeneity. Just as languages can make different choices as to the way they represent verbal properties in the temporal dimension (*Aktionsart*, *verbal aspect*), languages can also make different choices as to the way they represent nominal properties in the spatial dimension (*Seinsart*, *nominal aspect*). For instance, we can refer to the same thing as: ‘50 grapes’ (as when they are going to be distributed individually), ‘a pound of grapes’, or ‘a bunch of grapes’. In other words, in the act of referring different spatial features of the property ‘grapeness’ can be emphasized. It can be referred to as a number of distinct individual objects, as a mass, or as a collective entity (cf. Adams 1989: 3).

2.7. Conclusion

In this section I have proposed a classification of nominal subcategories in terms of the features Shape and Homogeneity and shown that across languages four of these noun types are used to refer to a singular discrete spatial entity (see note 1): singular object nouns (+Shape/-Homogeneity), set nouns (+Shape), sort nouns (-Shape/-Homogeneity) and general nouns (-Shape). I then argued that each nominal subcategory essentially defines a different mode of being or *Seinsart* of a spatial property, just as the various *Aktionsart* distinctions define different modes of action in the temporal dimension.

Notes:

(1) I am aware that noun phrases rather than nouns have referring potential; also that speakers rather than NPs are capable of referring; and that the immediate referent of an NP, i.e. the *discourse referent*, is a mental construct rather than an entity in the external world. But to avoid cumbersome formulations (such as: ‘a noun heading an NP that the speaker uses to refer to an entity in the discourse world and whose possible counterpart in the external world or Sein-correlate is a singular individual’), I will simply say from now on that a noun (of a particular subtype) or NP is used to refer to a certain entity—or similar statements to that effect.

(2) The term *Seinsart* was suggested to me by Simon Dik (personal communication).

(3) See also Quesada (1999: 218) on plural marking on the NP (rather than on the noun) in certain Chibchan languages; on noun phrase final number markers in the context of grammaticalization processes, see Lehmann (1995) and Frajzyngier (1997: 211–14).

(4) The plural marker used in this Nama example is the indefinite plural form; as the next example shows, Nama also has masculine and feminine plural forms.

(5) Notice, however, that according to Kastenholz (1989: 19) the unmarked Bambara noun is transnumeral in that it designates a concept (*Begriff*) rather than an individuated property (see section 2.3). To ‘individualize’ the noun the speaker must use an extra high tone on the second syllable (this is not possible in the case of high-tone monosyllabic nouns).

(6) On these so-called ‘plural words’, see also Dryer (1989c).

(7) See also Heath (1984: 153–4) on numerals in Nunggubuyu: ‘Numerals ... are also very often in predicative form when they are simple adjectival modifiers in English’; and (1984: 495): ‘Thus instead of saying “Two [men] went”, in Nunggubuyu it is more idiomatic to say “They were two, they went”.’

(8) This is only since traditionally Bantu noun classes have been defined as including number distinctions.

(9) According to Castrén (1858: 16) the plural marker in the noun is often omitted. Only NPs headed by an animate noun, however, agree in number with the verb. NPs headed by inanimate nouns, whether singular or plural, commonly have the verb in the singular.

(10) In complex numerals ($n > 5$) agreement patterns follow special rules; see Reh (1985: 254) for details.

(11) The tendency to mark number only in certain pragmatic contexts is also found in, for example, the Creole languages that are spoken in the Caribbean, where number is usually only indicated in definite NPs.

(12) At some point in his discussion of NP complementation, Cook (1984: 103) suggests that (basic) numerals are in fact relativized forms.

(13) Hoskison (1983: 51) writes about Gude: ‘With numerals greater than one, the singular form of the noun is usually used instead of the plural.’ On number marking in Chadic languages in general, see Zaborski (1986); Newman (1990); and Frajzyngier (1997: 196).

(14) Although Lango also has many nouns with a distinctive plural form (there is no regular plural formation), I have still entered Lango as a language with set nouns for two reasons. Even when a noun has a distinctive plural it is not always used and, secondly, it takes third person singular concord in the verb. Plural verb forms are only used with zero subjects and with the pronoun *gìn* ‘they’ (Noonan 1992: 167). More on this in Chapter 5.

(15) Although Hmong Njua is a classifier language, I will argue in Chapter 4 that it has set nouns rather than sort noun with the classifiers serving as nominal aspect markers, which will be introduced in Chapter 3.

(16) In Basque, however, the plural marker may be optionally present when the noun is modified by a numeral: *hiru neska(-k)* ‘three girl(-P1)A’ (‘three girls’). Indefinite NPs are not marked for number, which shows that number marking is at least partly determined by the fact that the plural marker has more or less amalgamated with the case/definiteness marker (more on this in Chapters 5 and 6).

(17) With human nouns, such as *ngoppa* ‘child(ren)’, the numeral can only be expressed in a predicative construction (van Baarda 1891: 19, 1908: 43); hence the presence of the pronoun *ja*. The NP *o ngoppa ja temidungi* (‘seven children’) may be glossed as ‘ART child(ren) they [are] seven’. In the 1908 edition van Baarda mentions that every once in a while a classifying word is added between a material noun and the numeral (van Baarda 1908: 44; on p. 22 of the 1891 edition he still calls these elements ‘samengestelde namen van hoeveelheid’ [composite names of quantity]). Since the so-called classifiers are not used very often (and seem to be restricted to natural materials and products, thus functioning as mensural classifiers rather than sortal classifiers), I have not taken such constructions into account. Furthermore the order Noun Classifier Numeral is extremely unusual from a cross-linguistic perspective (Rijkhoff 1990a) and the double occurrence of the ART(icle), which could also be regarded a noun phrase marker (Chapter 6), suggests that we might be dealing with an appositional structure, in which the numeral follows the measure noun (‘classifier’) in the second NP. In this view *o igo o liri saängo* [Art coconut Art three] ‘three bunches of coconuts’ could also be analyzed as [[Art coconut_N] [Art bunch_N three]].

(18) Gruzdeva seems to regard numerals in Nivkh as free elements, but recall that according to Mattissen and Drossard we are probably dealing with a polysynthetic construction. As to number marking she states (Gruzdeva 1998: 17): ‘A noun which is not a subject of a sentence (sometimes the one which is a subject) may have a singular form even though it refers to more than one referent.’ Earlier Austerlitz (1980: 241) wrote that ‘[in the South Sakhalin dialect of Gilyak] the plural morpheme is seldom, if ever expressed obligatorily in the noun’. Apparently ‘plural marking’ only more or less regularly appears with animate nouns in Nivkh, but never in the context of numerals and quantifiers (Mattissen and Drossard 1998: 14; Panfilov 1962–5: i. 191).

(19) The element glossed as ART(icle), ’o, also indicates that the referent is proximate and visible for the speaker.

(20) For a discussion of classifiers, see e.g. Foley (1997: 230–45) and contributions in Craig (1986b).

(21) Nivkh has no proper sort nouns, since classifiers are only used with numerals up to five; higher numerals appear without a classifier (see Mattissen and Drossard 1998: 21 for references regarding this topic); see also notes 18 and 20.

(22) It is sometimes said that Turkish also has a classifier, namely *tane* (Lewis 1967: 80–1; Underhill 1976: 126–9). Schroeder (1999: 95–109), who incidentally found that *tane* is not discussed in any modern Turkish grammar, investigated the use of this so-called classifier and found that it never appears with numbers over 20, in NPs containing a determiner (i.e. in NPs referring to topical or identifiable referents), and in

NPs with accusative case (Schroeder 1999:103–4). He concludes (1999: 108) that the use of *tane* is not determined by the lexical semantics of the head noun, but rather by pragmatic considerations (i.e. ‘by the question as to what role a certain referent is to play in the discourse’).

(23) Samoan and Quechua were also discussed in Chapter 1. As to the flexible nature of the nouns/adjectives in Hurrian and Turkish, see Speiser (1941: 72) and Lewis (1967: 53 f.) respectively (on Turkic languages see also e.g. Menges 1968: 107).

(24) There is some doubt whether the plural marker is really obligatory without a numeral in Quechua. Perhaps the plural marking is an innovation, possibly a borrowing, since in some Quechua dialects and languages *-kuna* is absent (see Simon van de Kerke 1996 and personal communication; Lefebvre and Muysken 1988: 108).

(25) Which of these two phonemic variations of the so-called plural morpheme in Turkish (*-lerf-lar*) is used depends on the palatality of the preceding vowel. Turkish also displays another feature associated with set nouns, namely *number discord* (see section 4.2.1.2.), in that ‘a singular verb is commonly used with an inanimate plural subject’ (Lewis 1967: 25–6, 246 f.); for a detailed discussion of number discord in Turkish, see Schroeder (1999: 111–25). Note furthermore that the so-called indefinite article *bir* (at least when combined with a noun used to refer to a discrete spatial object) can be analyzed as a singulative aspect marker (see section 10.2.1.2.4).

(26) Underhill (1976: 92–3) writes: ‘In the third person plural *-ler/-lar* may precede the third person *-(s)I(n)*, giving the combination *-lerI(n)/-lAri(n)*. As we have seen before, *-ler/-lar* is used only when there is no other indication of plurality of the possessor: *oda-lar-I* [room-PL-3:POS.SX] “their room”. If there is an explicitly plural possessor, *-ler/-lar* is omitted: *cocuk-lar-In oda-SI* [child-Pl-3:POS.SX room-3:POS.SX] “the children’s room”. The possessed noun may itself be plural: *cocuk-lar-In oda-lar-I* [child-Pl-3:POS.SX room-PL-3:POS.SX] “the children’s rooms”. Even if both the possessor and the possessed nouns are plural, only one *-ler/-lar* is permitted; thus *oda-lar-I* [room-PL-3:POS.SX] can mean “his rooms”, “their room”, or “their rooms”.’

(27) I have no information about ‘number’ marking on nouns and verbs in the two other languages with flexible predicates (Samoan and Hurrian), but note that the same phenomenon is attested in Fijian (which is closely related to Samoan and which also seems to have flexible predicates). Churchward (1941: 16) wrote: ‘It would appear, indeed, that the original or fundamental meaning of *vei-*, with verbs as with nouns ... is just that of plurality, or rather collectivity (a number of things forming a collection or a group).’ Consider also the following statement by Dixon (1988: 175) in his grammar of Boumaa Fijian: ‘*Veि-* can be prefixed to some nouns and time words, most kin terms ..., and to many verbs (sometimes with and sometimes without the passive suffix); in each instance it has a collective sense.’ See also Mithun (1988) and Newman (1990) on verbal and nominal plurality.

(28) From a cross-linguistic perspective obligatory number (plural) marking seems to be the exception rather than the rule, since in many, perhaps even most languages, number distinctions in the NP are on the whole considerably less relevant, i.e. either optional or altogether absent. Cf. Boas (1911a: 37–8) and Mithun (1999: 82) on Amerindian languages, Wurm (1982: 90) on Papuan languages; Jones (1970) and Bisang (1993, 1996) on SE Asian languages; Dixon (1980: 22) on Australian languages.

(29) Without the numeral the noun may optionally be provided with a so-called plural affix when the speaker of Oromo means to refer to more than one individual; this is discussed in more detail in section 4.2.1.2.

(30) Interestingly, the classifier is not always used with Spanish numerals (Yucatec is spoken in Mexico), which are often used with referents involving four or more units or individuals. However, since ‘the syntactic significance and interpretation of these mixed constructions is quite complicated’ they are not discussed in Lucy’s study (Lucy 1992: 43, 51).

(31) See Hundius and Kölver (1983: 167 f.) for differences between sortal and mensural classifiers in Thai (cf. also Adams 1989: 2–10).

(32) On mass nouns, see e.g. Bunt (1985); Drossard (1982); Gil (1987).

(33) Data from a cross-linguistic study on genericness by Claudia Gerstner-Link (forthcoming) suggest that there is another interesting difference between singular object nouns and set nouns. When a set noun is used in a generic NP one always finds the morphologically unmarked (transnumeral) form. This is not the case with singular object nouns, which may also appear with a plural marker.

(34) See Burling (1965); Greenberg (1972); T’sou (1976); Lyons (1977: 227, 462); Adams (1989: 3–9).

(35) The full quotation reads: ‘Interpretatively, in Yucatec all nouns are like our mass nouns in that they are neutral with respect to logical unit or shape.’ Note, however, that Yucatec nouns are not classified as mass nouns in the classification of nominal subcategories presented here. Notice also that some Yucatec nouns that are used with certain animate entities are perhaps better categorized as *set nouns* in that they may occur with an optional so-called ‘plural marker’ which is absent when the NP contains a numeral (see also Chapter 4): ‘Yucatec lexical nouns can be marked for plural by suffixing -ó’ob to the lexical head noun, so pèek’ “dog”, plus -ó’ob “[plural marker]” yields pèek’-ó’ob “dogs”. This suffix is identical to the one suffixed to verbs to indicate the presence of third person plural complement. The suffix is usually used to indicate explicitly that the noun phrase refers to a multiplicity of referents; sometimes it apparently suggests that all of a given group of referents are being referred to.

The suffix is optional or *facultative* in that it need not be used for correct reference when a multiplicity of referents does in fact exist, but it can be used to clarify or emphasize such multiplicity. For example, *yàan pèek’ té’ elo’* “there-is-dog over-there” is ambiguous syntactically as to whether it refers to a multiplicity of referents or not. Specifically, it can refer to more than one dog: “there are dogs over there” (Lucy 1992: 46 f., 49). Furthermore this ‘plural’ marker is properly speaking not an inflectional element (like the English plural suffix -(e)s) but rather a bound pronoun which is formally identical to the third person plural person affix and which can also be interpreted as a possessive marker (Lucy 1992:47).

(36) The terminology chosen here represents an improvement over the labels chosen previously (e.g. Rijkhoff 1991) in that now the label of a subcategory reflects the conceptual characteristics of the denotata of the noun belonging to this subcategory.

(37) Recall that this is a classification of basic (underived, uninflected) nouns. Hence it does not apply to e.g. plural forms of singular object nouns, which also help to define agglomerative entities: ‘apple(s)’ added to ‘apples’ are also ‘apples’.

(38) Note that in languages with sort nouns and general nouns the notion Shape always plays a prominent role in the classifier system and that Shape is essential in the process of localization and identification: one can only tell where one object stops and another begins when there is a spatial boundary that separates them.

(39) For the same reason *Seinsarten* (section 2.6) are not quite the same as *modi essendi*, which were developed by a group of medieval grammarians known as *Modistae* and ‘which were meant to provide an ontological foundation for grammar’ (Gabler 1991: 567). Although both terms can be translated as ‘modes of existence’ or ‘modes of being’, *modus essendi* is an *ontological* notion and concerns the thing itself. *Seinsart* on the other hand relates to the property as denoted by the noun (esp. how it is specified in terms of the features Shape and Homogeneity) and is thus a linguistic category. Different languages may use different nominal subcategories (*Seinsarten*) in connection with the same thing in the extra-linguistic world. For example, whereas the real-world entity ‘table’ is a discrete physical object, the referent of an NP headed by the equivalent of the noun ‘table’ in Language L may be (depending on the nominal subcategory employed by speakers of Language L) a general entity, a sort entity, a set entity, or a singular object entity.

This does not necessarily mean that the speakers of e.g. Thai or Yucatec Maya do not *know* that a table in the physical world is a discrete object, but rather that this particular piece of knowledge is simply not part of the lexical semantics of the noun (instead it would be part of the encyclopedic knowledge about the referent of the NP; for a similar point see e.g. Unterbeck 1993). However, the idea that the employment of different nominal subcategories in varying languages for the same object *does* reflect ‘differences in the ontological beliefs the speakers of these languages hold about the referents of nouns’ is discussed in e.g. Foley (1997: 231; on the subject of *ontological relativity* see also Quine 1960, 1969; Lucy 1992; Imai and Gentner 1993).

(40) There is nevertheless a connection in that e.g. nouns with the feature [+Homogeneity], i.e. mass nouns and collective nouns, are normally not used to refer to entities whose *Sein*-correlate is a single, discrete spatial object.

(41) The same goes for the referent of a clause, i.e. a temporal entity (see Chapter 7).

(42) See for example Dik (1980a: 38 f.); Mithun (1984: 950 f.); de Reuse (1994); Cook and Wilhelm (1998).

(43) It depends on the time span between T_i and T_j whether the property or relation designated by the verb can be further characterized as momentaneous or durative.

(44) The semantic subcategorization of verbs in terms of *Aktionsarten* also involves such features as Dynamicity and Control (see e.g. Dik 1997: 115), but for the sake of the argument I have restricted myself here to the features that also play a role in the verbal aspects.

3 Nouns: Real and Apparent Nominal Subclasses

3.1. Introduction

On the basis of various criteria nouns in many languages can be divided into different classes, including genders, which co-determine the proper expression (and meaning) of linguistic structures in terms of form, order, or tone. Thus, one can find references to, for example, grammatical (i.e. lexico-semantic), syntactic, morphological, or phonological noun class systems in the linguistic literature. Despite the vast amount of literature on gender and noun class systems, our knowledge of the impact that such systems can have on linguistic expressions still leaves much to be desired.¹ Therefore this chapter cannot offer much more than a rather arbitrary and superficial (but nevertheless useful) collection of data that pertain to noun class and gender systems that are somehow relevant for the proper expression (morphology, syntax, phonology) of NPs and sentences.

It must be mentioned at the outset that one of the major problems in the area of noun classification is that it is often difficult to determine whether one is dealing with a property of the noun as it occurs in the mental lexicon (i.e. with a real noun class system) or with a property of the ontological correlate of the referent of the NP in the physical world (henceforth in this chapter simply ‘referent’). As we will also see below, this holds especially for classifications which involve such features as <Human> or <Animate>; strictly speaking one is only dealing with a real noun class system if the classification is due to some property of the noun.² For instance, certain languages consistently indicate the sex of a human or animate referent in some way or other, in which case the phrase ‘natural gender’ (p.61) (or ‘semantic’, ‘biological’, or ‘sex gender’) is used, as opposed to ‘grammatical gender’ (i.e. based on a lexical feature of the noun, as in e.g. Dutch and German). If there is a perfect correlation between the sex of the referent and the natural gender class of the noun (animate masculine, animate feminine, neuter), it is impossible to decide whether in the language user’s mind the classification is based on a lexical feature of the noun, or on a property of the referent of the NP, or both. In some languages this sex distinction (male animate, female animate) also includes inanimate nouns (see below). Although this may seem strange from one point of view, the division of certain inanimate nouns into the masculine and feminine gender may be perfectly logical in a culture where inanimate entities are conceived as male and female mythological figures (Dixon 1986; Lakoff 1986, 1987; Harriehausen 1990: 97).

In other languages it seems that the natural genders have been metaphorically extended, so that, for example, everything big, strong, longish, etc. is classified as masculine; here the biological foundation is rather obvious. Again, in cases such as these it is rather difficult to decide whether or not the classification is wholly or partly due to a feature of the noun. However, there are also clear-cut cases such as Spanish, where all nouns are classified into two genders (masculine, feminine), which often bear no obvious relation to properties of the referent (but cf. Zubin and Köpcke 1986 on German).

In any case, nouns in many languages are often classified on the basis of some common semantic feature, e.g. <Human/Animate>, <Masculine>, <Feminine>, <Round>, <Long>, <Edible>, <Dangerous>, etc.³ The following example is from Fula (Gombe dialect; Niger-Kordofanian); the classificatory elements are in bold print:

Fula (Arnott 1970: 12)

(1) sawru	mawndu	ndu'u
stick	big	this
'this big stick'		

In Gombe Fula nouns are classified into twenty-five different classes. Thus *sawru* and other nouns of its class designate circular and cylindrical objects. Each noun class correlates with different agreement patterns, compare:

Fula (Arnott 1970: 75)

(2) wudere	mawnde	nde'e
cloth	big	this
'this big cloth'		

(p.62) These examples show how noun class determines the *form* of a constituent. The next examples from MalakMalak (Australian) illustrate that noun class can also determine the *order* of constituents. In this language inalienable nouns (i.e. nouns which always require reference to a possessing entity) must follow, whereas alienably possessed nouns must precede the possessor noun (here *yin'y'a* 'man'). Notice also that in the case of alienable possession -*nö* is suffixed to the possessor noun.⁴

MalakMalak (Birk 1976: 106, 107)

(3) yin'y'a	puntu	[inalienable]
man	head	
'man's head'		
(4) muyin'y	yin'y'a-nö	[alienable]
dog	man-NÖ	
'man's dog'		

Reflexes of noun class distinctions may also occur outside the NP. In the following example from Swahili (Niger-Kordofanian) noun class is also indicated in the (non-verbal) predicate of the main clause (*vikali* 'sharp'):

Swahili (Allan 1977: 286)

(5) Visu	vidogo	viwili	hivi	ambavyo
----------	--------	--------	------	---------

knife	small	two	this	which
nilivinunua	ni	vikali	sana	
bought: 1SG	be	sharp	very	

'These two small knives which I bought are very sharp'

This chapter gives examples of both real and apparent noun class systems, but to the extent that this is possible I will attempt to distinguish between them. For ease of reference, however, I will generally continue to use the phrase 'noun class' or similar terminology, even if it turns out that we are actually dealing with an apparent noun class system, i.e. when the classification is based on some feature of the referent (or, rather, its ontological counterpart in the physical world). First I will devote some attention to real and apparent noun class distinctions that relate to the expression of constituents outside the domain of the NP (such as the form of the main predicate), then I will concentrate on formal consequences of (real and apparent) noun class distinctions with respect to constituents in the NP proper.

I will not go into the distinction between mass nouns, singular object nouns, etc.; this has already been discussed in Chapter 2 on nominal subcategories or *Seinsarten* and concerns the way properties designated by nouns are represented in mental space in terms of the features Shape and Homogeneity.

(p.63) 3.2. The reflection of real and apparent nominal subclass distinctions outside the NP

This section discusses certain correlations between nouns or referents of NPs on the one hand and constituents *outside* the NP such as relators (adpositions, case affixes) and the main predicate on the other.

3.2.1. Positional verbs

3.2.1.1. Existential and locative constructions

In at least four languages in the sample nouns can be classified according to the predicate (a so-called *positional verb*) with which they occur in an existential construction (see also Hengeveld 1992b: chapter 10): Dutch (Indo-Hittite), Ika (Amerind, Chibchan), Koasati (Amerind, Penutian), and Sarcee (Na-Dene). Outside the sample the same phenomenon is also attested in the Siouan and Papuan languages (e.g. Asmat, Enga, Kombai), where it is often regarded as some kind of covert form of noun classification.⁵ Thus, in Asmat (Drabbe 1959: 16)

the 'positional' verbs have various functions. In the first place, their natural function is to indicate expressly a certain position in which the subject finds itself, but then also to express simply a presence, similar to our 'being there', which in passing also indicates the position of the subject. According to usage then, they speak of certain classes of things that they stand or sit, etc.; a snake or a crocodile lies, a bird sits on a branch or on the ground, but hovers in a nest or in space. A strange object 'hovers' in the human body; a man is understood to sit, especially to stand, a house is located; *se-* is used for rivers, for water in the river or the sea, and for everything situated on or in the water.

From this it appears that it is usually the perceived position or shape of the referent that

is at stake here, rather than a property of the noun. Although ‘man’ usually occurs with ‘stand’ and ‘woman’ with ‘sit’ etc., the classification, if one may call it that, is not rigid. For instance, in Enga a standing tree typically collocates with ‘stand’, but a tree that has been cut down will occur with ‘lie’.

Enga (Foley 1986: 90)

(6) ítá	dúpa	kate- <i>ngé</i>
tree	the	stand-HAB
‘there are trees’		
(7) ítá	poká-pae	sí- <i>nge</i>
tree	cut-STATE	lie-HAB
‘there are (felled) trees’		

(p.64) In Koasati, too, positional verbs are used in existential constructions, where the choice of the main predicate is correlated with the shape of the referent. In this language the verbs have a singular and a plural form. The latter is often not the regular plural, but a suppletive form taken from another positional verb (Kimball 1991: 452–61; see also Cook 1984: 55 on Sarcee).

Existential and locative clauses in Ika make use of a ‘noun class indicator plus copula to indicate existence or location’ (Frank 1990: 53). Referents are characterized as: one-dimensional (long) objects, two-dimensional (flat) objects, three-dimensional objects, liquids, containers, or objects with specialized holders, cf.:

Ika (Frank 1990: 53)

(8) tšo?kui-se?	džé	a?kuaskuasi	zina
gourd_bowl-LOC	water	LIQUID	COP
‘The water is in the bowl’ or ‘There is water in the bowl’			

Ika verbs of placing are also sensitive to dimensional and other physical properties of the referent. For example, the different forms of the equivalent of ‘put down’ vary according to the characteristics and position of the referent:

Ika (Frank 1990: 55)

(9) kʌn	gakó	ú
stick	put_down:LONG	AUX
‘Put down the stick’		
(10) ribru	pa	ú
book	put_do wn: FLAT	AUX
‘Put down the book’		

In the case of Ika too it is clear that we are not dealing with noun classes despite the fact that this is treated as such in the grammar (Frank 1990: 56):

The noun class ‘upright’ refers to the position or existence of items such as pots, plates, people, or plants that have a vertical orientation. Note that people and plants only fall into this class when in an upright position. Thus, when approaching a person who is standing, a polite comment is *ei tšo* (thus upright) ‘you’re standing’. A person is classified as three dimensional (*sa*), however, when sitting.

To some extent similar phenomena are found in Dutch. For example, different verbs of ‘putting’ are used in relation to different kinds of objects:⁶

Dutch

(11) Leg dat boek neer!

lay that book down

‘Put down that book!’

(p.65)

(12) Zet die tas neer!

Put that bag down

‘Put down that bag!’

The verb (*neer*) *leggen* ‘to lay (down)’ is used with flat objects, whereas (*neer*) *zetten* ‘to put/place down’ is typically used with three-dimensional objects. Furthermore, positional verbs are used in locative/existential sentences:

Dutch

(13) Er was/stond een man

there be/stand:SG.PAST a man

‘There was a man (standing)’

(14) Er lag een boek

there lie:SG.PAST a book

(op (de)tafel)

(on (the)table)

‘There was a book (lying on the table)’

(15) Er stond een vaas (op (de) tafel)

there stand:SG.PAST a vase (on (the) table)

‘There was a vase (standing on the table)’

(16) Er zat een vlek (op mijn jas)

there sit:SG.PAST a spot (on my coat)

‘There was a spot (on my coat)’

Cross-linguistically positional verbs such as ‘sit’, ‘stand’, and ‘lie’ tend to grammaticalize in two directions. Either they become copulas (illustrated by the examples from Asmat, Enga, and Dutch above; cf. Devitt 1994) or they develop into tense or aspect markers (cf. Bybee et al. 1994: 129–31). The latter development is also attested in the Dutch progressive construction, as is shown in the following example:⁷

Dutch

- (17) Hij zat/lag/stond de krant te lezen
 He sit/lie/stand:PAST.SG the newspaper to read
 ‘He was reading the newspaper’

It is clear that the Dutch positional verbs have not quite lost their positional meaning in their derived usage, but in the case of e.g. Spanish *estar* (‘to be’), which originates from Latin *stare* ‘to stand’, the verb can be used with objects in any position (Stengaard 1991; cf. also Dutch *staan* ‘to stand’ and *be-staan* ‘to exist’).

(p.66) 3.2.1.2. Possessive constructions

The present section concentrates on Onondaga (Amerind). This language has a particularly interesting noun class system (but see below), which is relevant with respect to both incorporation and possessive constructions. If possession is expressed sententially the ‘possessed noun’ is incorporated into a positional verb. Two subclasses of nouns are distinguished in relation to this type of possessive construction: ‘nouns designating movable objects and nouns designating fixed objects. Movable objects are incorporated into the positional verb *-yv-* “be lying” to indicate possession.’ Witness:

Onondaga (Woodbury 1975b: 39)

- (18) hoheská:yv?
 it:him.arrow.be_lying:ASP
 ‘He has an arrow’

When fixed entities are involved, which include body-parts in Onondaga, the choice of the incorporating verb varies according to the shape or position of the referent; compare (1975b: 40–1):

Onondaga (Woodbury 1975b: 40)

- (19) hothu:tó:ta?
 it/him.REFL.tree.be_standing_upright:ASP
 ‘He has a tree’
- (20) honu?waú:ta?
 it/him.head.be_projecting_outward:ASP
 ‘He has a head’

It is not entirely clear if a fixed noun always collocates with the same positional verb in this construction. What is clear is that the classification is not so much based on some feature of the noun, but rather on some property of the referent.

3.2.1.3. Noun-incorporation

Besides the classification that is made on the basis of the kind of positional verb with which a noun will occur in a possessive construction, there are two additional parameters along which nouns are classified in Onondaga: (i) whether a noun, which superficially heads an NP that has Patient (or Goal) function, is always, sometimes, or never incorporated; and, if it is incorporated, (ii) with which positional verb the noun must collocate.

For a better understanding of (so-called) noun class systems in Onondaga, I first present some facts about verb classes in this language. Onondaga has a group of verbs that incorporate obligatorily (e.g. adjectival predicates and positional verbs), those that incorporate optionally (the largest group), and those that cannot incorporate (e.g. what Woodbury (1975b: 44) calls ‘secondary contact verbs’, (p.67) verbs that do not involve direct contact between agent and patient, such as ‘weigh’, ‘follow’, ‘look at’).

As for the nominals in Onondaga, there are nouns (heading Patient NPs denoting inanimate fixed objects) that, if possible, are always incorporated in the verbal complex; there are other nouns that may be incorporated, and yet another group is never incorporated (kinship nouns and noun roots describing adult humans) or can only be incorporated if provided with a nominalizer (Woodbury 1975b: 45–53; see also Woodbury 1975a: 15).⁸ The predicate ‘house’ belongs to the first group in that it usually does not occur in a separate NP (1975b: 47):

Onondaga (Woodbury 1975b: 48)

(21) Harry wa?hanuhsahní:nu?

Harry TNS/he:it_house_buy:ASP

‘Harry bought a house’

Even when fixed nouns like ‘house’ or ‘hill’ must be expressed as part of a distinct phrase (for instance, because the main verb belongs to the group of verbs that cannot incorporate), they are still incorporated in a positional verb:

Onondaga (Woodbury 1975a: 15)

(22) Harry wa?hatkathwá?

kanuhsá:yε?

Harry TNS.he/it.look_at:ASP it.house.be_lying:ASP

‘Harry looked at a house’

However, here we must distinguish between two kinds of fixed nouns. On the one hand there are the truly fixed nouns such as ‘hill’, ‘earth’, ‘island’, and body-part nouns, which ‘select a single positional verb which classifies them’. (Woodbury 1975b: 45):

Onondaga (Woodbury 1975b: 43)

(23) onu:tó:ta?

it.hill.be_standing_upright:ASP

'hill'

(24) ohwv:tsyá:te?

it.earth.exist_in_a_neutral_position:ASP

'earth'

On the other hand there are nouns that 'describe objects that are hard to move or which are habitually kept in a given location' (ibid.) such as 'kettle', which can occur with a different positional verb at different times (ibid. 46). Again the classification seems to be based on properties of the referent.

(p.68) Onondaga (Woodbury 1975b: 46)

(25) kana?tsyaniyú:ta?

it.kettle.be_hanging.ASP

'The kettle hangs'

(26) vkana?tsyótak

TNS.it.kettle.be_standing:CONT

'The kettle will stand'

3.2.2. Classificatory verbs

Properties of entities may also correlate with verbs in non-existential/locative and non-possessive constructions, in which case we may speak of *classificatory verbs*. It is not quite clear whether this always involves a basic verbal predicate or whether we are dealing with a classificatory element that is affixed on or incorporated in the predicate (cf. Allan 1977: 287; Dixon 1982: 223; Barron 1980: 12 f.). In any case it seems that there is no productive process of classifier affixing or noun-incorporation (Blankenship 1997: 98); apparently languages only have a restricted number of classificatory verbs. In the sample such verbs are attested in Sarcee (Na-Dene). Outside the sample, classificatory verbs are attested in other languages of the Na-Dene phylum, but also in languages of the Amerind phylum.⁹ In Sarcee, according to Cook (1984: 54–5; 139 f.),

there are strict selectional restrictions between the subject and/or object noun and the predicate verb. For instance, in an English sentence such as 'I picked it up,' the verb is relatively free of selectional restrictions since the class of nouns that may co-occur as objects (i.e. nouns that it refers to) is large, not limited to a small subclass. Corresponding to this single sentence, there may be several Sarcee sentences, depending on the class of nouns involved, for example,

) a. -?ó(n-) nádísís?ó 'I picked it up' (e.g. a rock)

- b. -tó(n-) nádísís?ó ‘I picked it up’ (e.g. a stick)
-
- c. -tíh nádísístíh ‘I picked it up’ (e.g. a child)
-
- d. -tcùz nádísísstsùz ‘I picked it up’ (e.g. a quilt)
-
- e. -ló nádísísstló ‘I picked them up’ (plural inanimate object).

In the case of Sarcee classificatory verbs, then, it seems that it is the properties of verbal predicates that are at issue. Apparently these verbs (Sarcee has eight sets of them) have acquired a highly specified meaning, which is reflected in the selection restrictions of the arguments. In this respect they are not basically different from other verbs or any other kind of predicate for that matter, since generally speaking any selection restriction imposes a classification on terms (p.69) (and their head nouns) to be inserted into the argument positions (for instance, the predicate ‘give’ requires an animate Agent and Recipient, and an inanimate Goal or Patient; cf. Grinevald 2000: 68). The various forms expressing the relation ‘to pick up’ are only different to the extent that they have slightly different selection restrictions for the second argument (the Goal or Patient). The relevant features are: <+Animate>, <+Round>, <+Long>, etc.

3.2.3. Classificatory elements in the verb stem

There are languages, such as those from the Waris group (Indo-Pacific), which employ classificatory prefixes that are said to ‘specify salient semantic features of their object noun’:

Waris (Foley 1986: 90, 91)

- (28) wonda ka-m mwan-vra-ho-o
net bag I-DAT CLF-get-BEN-IMP
'Give me a net bag'
-
- (29) nenas ka-m li-ra-ho-o
pineapple I-DAT CLF-get-BEN-IMP
'Give me a pineapple'

Mwan- is the classifier for soft pliable objects like net bags, skirts, bark mats, while *li-* denotes oblong fruit objects like pineapple, ears of corn or pandanus fruits’ (Foley 1986: 90–1). However, since ‘multiple class membership is not only possible, but the rule’ (W. Seiler 1986: 201), it is doubtful whether we are dealing with a true example of a noun class system here.

Whereas classificatory prefixes in Waris probably derive from verbs (W. Seiler 1986), in Caddo (Amerind) the classificatory element in the verb originated as an incorporated noun.¹⁰

Caddo (Mithun 1986: 386)

(30) Kapí: kan-čâ:ni ‘ah

coffee LIQUID-buy:PAST

‘He bought (liquid) coffee’

(31) Kapí: dân:-čâni’ah

coffee POWDER-buy:PAST

‘He bought (ground) coffee’

Mithun (*ibid.*) argues that the classificatory element does not reflect properties of the noun, but rather characteristic features of the referent. She also suggests that (p.70) once the process of incorporation has stopped, the incorporated nominals turned classificatory elements may lose their transparency. This may lead to ‘an ever diminishing pool of increasingly opaque relic compounds’ (*ibid.* 392). It can be hypothesized that eventually the verb-plus-classificatory element ends up as a distinct verbal predicate not unlike the kind of classificatory verb mentioned above in connection with Sarcee (cf. also Mithun 1984).

3.2.4. Predicative numerals

There are languages in which cardinality may or must be expressed predicatively with certain nouns. In Galela (Indo-Pacific), for example, a predicative numeral must be employed in relation to human nouns (cf. section 5.2.2):

Galela (van Baarda 1891: 19; also van Baarda 1908: 43)

(32) o ngoppēdikka ja sinotto

ART woman they two

‘two women’ (lit. woman they [are] two)

In Yurok (Uralic-Yukaghir) predicative numerals contain in addition a classificatory element. Since it seems that at least some nouns may occur with different classificatory elements it is possible that extra-linguistic factors are involved here (see for references Dixon 1982: 224).

3.2.5. Relators and participant coding

In many languages a basic distinction is made between human/animate and non-human/non-animate nouns or referents, which is relevant both inside and outside the domain of the NP proper (Smith-Stark 1974; Comrie 1980 and 1981b: chapter 9; Dik 1997: 34–6). Outside the NP the significance of this division is manifested in the relator system (adpositions, case affixes) and in the coding of participants in or on the verbal complex. Recall that in some languages (such as Mangarayi, see below) the class of humans has been extended to non-humans (cf. also Meillet 1964).

3.2.5.1. Role marking

In some languages the form or occurrence of a relator (preposition, postposition, case affix) is dependent on whether or not humans or animates are involved. This is the case

in at least the following languages in the sample: Basque, Ket, Korean, Krongo, Hittite, Sumerian, and Tamil.

In Basque some relators, such as those marking the locative, the ablative, and the allative, have variant forms for animates and inanimates (Saltarelli 1988: 200).

Ket has three genders (or classes), masculine, feminine, and neuter, which to a large extent correlate with properties of the referent (i.e. sex distinctions: male animate, female animate, and neuter). However, there are some exceptions. For (p.71) instance, most fish names and words for large wooden objects have masculine gender, whereas small wooden objects are feminine. The gender/class distinctions are manifested in, for instance, the spatial cases, which have slightly different suffixes for different classes (allative *ob-danya* ‘to father’ and *ab-diya* ‘to mother’). Etymologically these suffixes may be related to pronominal forms (Comrie 1981a: 263), which suggests that originally referents were cross-referenced on the relator and that at some point in the evolution of Ket these pronominal elements became integrated in the relator (cf. de Groot and Limburg 1986).

Korean has two allomorphs meaning ‘to, toward’: *ege* and *e*. The former collocates with animates, the latter with inanimates (Lee 1989: 56).

Krongo has two variants of both the locative and the ablative case prefix. One is used with humans and the other with inanimates. In the case of animals either one may be used, the selection being dependent on the degree of affection one has for the animal in question (Reh 1985: 146–7, 149–50).

Mangarayi (which is not in the sample) provides a slightly variant case. It has three genders: feminine, masculine, and neuter. The masculine and feminine genders include not just all female and male nouns, but also, e.g., some kinship terms, higher animals, and mythological beings. The only other gender, the neuter, constitutes a residual category. Interestingly, the masculine and feminine nouns take nominative-accusative case marking, while the neuter nouns pattern ergatively-absolutively (Merlan 1982: 56; for an explanation see Silverstein 1976; cf. also Du Bois 1987; see Garrett 1990 on the origin of NP split ergativity and Siewierska 1996b on different alignment systems). Similarly, in Hittite neuter nouns inflect ergatively and common gender nouns (comprising masculine and feminine nouns) inflect accusatively (Garrett 1990).

In Sumerian humans occur with the dative, whereas non-humans are combined with e.g. the locative and the ablative case (Thomsen 1984: 49, 88).

Asher (1982: 107, 136–7), in his description of Tamil, states that in some dialects animate object NPs must carry the accusative suffix, whereas for inanimates this suffix is sometimes optional.

Some of these examples indicate that there is a tendency for classificatory elements to cluster with locative elements. This has also been recognized by Allan (1977: 287–8), who distinguishes a type of language that he calls intra-locative classifier languages, in which classificatory elements are ‘embedded in some of the locative expressions which obligatorily accompany nouns in most environments’. For instance, Toba (Amerind) has three prefixes which occur in NPs denoting visible referents. These prefixes correlate with vertical (extended) entities, horizontal (extended) entities, and saliently three-dimensional entities.

Similarly, Eskimo (Eskimo-Aleut) seems to use different portmanteau morphemes in relation to visible referents, which express both the relative location of the real-world counterpart of a referent in the physical world ('here', 'there', 'up there', 'down there') and the notion ±extended (saliently long).

(p.72) Finally, in some languages the occurrence of a relator in a possessive construction depends on whether the noun/referent requires reference to a possessing entity in that it only appears in the case of an alienably possessed noun (see section 3.3.2.3 on the difference between alienable and inalienable possession). This is, for instance, attested in Bambara and Lango:

Bambara (Bird 1972: 6, 8)

(33) Bàba	ba	
Baba	mother	
'Baba's mother' [inalienable]		
(34) Bàba	ka	so
Baba	GEN	house
'Baba's house' [alienable]		

Lango (Noonan 1992: 157)

(35) pyèn	ájáŋâ	
skin	cat	
'the cat's skin' [inalienable]		
(36) pyènn	à	dákô
skin	AP	woman
'the hide belonging to the woman' [alienable]		

3.2.5.2. Participant coding

In quite a few languages only certain kinds of referents are explicitly coded in the verbal complex. In such cases the language usually distinguishes between (third person) human/animate and non-human/inanimate entities. For example, in Hixkaryana, the particle *komo* 'collectivity' is usually only (optionally) expressed in the verbal complex in the case of humans (Derbyshire 1979: 127, 136, 145–6; see also 3.3.3.1 below).

3.2.6. Honorific nouns and verbs

Korean is one of the languages that has so-called honorific nouns, which signal respect to the referent of the NP. As a rule an honorific noun (each of which can be paired with a corresponding plain noun) is used with an honorific verb (if it is available), but the correlation is not strict (Lee 1989: 57–8). For example, 'father' can be translated as *abʌji* (plain) or *abʌnim* (honorific). Only the latter form usually occurs with an

honorific verb, but if no such verb is available the plain verb will carry the suffix *-si* or *-si*. Apparently, honorific and non-honorific nouns constitute true nominal subclasses, because the choice of the verb depends on the noun inherent feature <±Honorific>. On the other hand, the decision to use an honorific noun, rather than its plain counterpart, is motivated by extra-linguistic factors, such as socio-cultural conventions (Rijkhoff 1995, 1998).

(p.73) 3.3. The reflection of real and apparent nominal subclass distinctions inside the NP

In this section I describe some of the ways real and apparent nominal subclass distinctions are formally reflected *inside* the NP. Dixon (1986) has argued that there are two kinds of (semantically based) classificatory systems that pertain to nouns: noun class and noun classification systems (including numeral classifiers), each of which has its own set of formal characteristics (see Table 3.1). It must be emphasized that these sets of characteristics only relate to prototypes; in practice it is often rather difficult to distinguish between the two systems. In the case of Nasioi, for example, it is explicitly stated that the classification system is of the ‘intermediate type, not corresponding neatly either to the numerical classification systems typified by southeast Asian languages, nor to the more usual noun class/gender languages of New Guinea, but showing features of both’ (Foley 1986: 84–5; see also Rausch 1912; Hurd and Hurd 1966; Hurd 1977).¹¹ This language has over a hundred semantically very specific classificatory suffixes which are used with nouns, adjectives, numerals, demonstratives, and nominalized verbs. Still, to the extent that the two systems can be kept separate, it has been argued that historically noun class systems may have derived from numeral classifier systems, although the evidence to support this claim is still rather poor (Corbett 1991: 137f., 310f.).

Greenberg (1978a) has argued that nouns may acquire gender markers through the affixation of an erstwhile classified demonstrative (the same process may apply to unclassified demonstratives, in which case they will simply end up as noun markers). He recognizes four stages in this diachronic process (which he illustrates with many synchronic examples from various languages in different stages of their development; more on this in section 3.3.2.4.1). The question is, of course, how this demonstrative got to be marked as a classificatory element in the first place. As a possible answer, Greenberg suggests that numeral classifiers might be responsible for this. He had observed earlier (Greenberg 1972) that the demonstrative pronoun is the first constituent with which the numeral classifier occurs if it combines with other constituents than the numeral (see also Dixon 1982: 170 f., 230 f.).¹²

(p.74)

Table 3.1. Properties associated with prototypical exemplars of classifier and noun class systems

Noun (incl. numeral) classifier	Noun class
Lexico-syntactic phenomenon	Grammatical phenomenon
Fair number of classifiers	Smallish number of classes
Does not involve all nouns	Involves all nouns

Classificatory element is a free form (except with numeral), i.e. it is never morphologically fused with the noun	Classificatory element is expressed by inflectional morphology, separate grammatical words, clitics, affixes
Classificatory element never outside NP	Classificatory element always on other constituent (besides N) and may also occur outside NP
More variation; classifier use often indicative of style/mode differences	Little variation between speakers (because obligatory morphological system)

Source: Based on Dixon (1986).

In the sections below I will first discuss cases of noun classification, which includes both noun classifiers (3.3.1.1) and numeral classifiers (3.3.1.2), then I will deal with noun class systems (including genders; section 3.3.2).¹³ The section on noun classes is not only concerned with the traditional grammatical (lexico-semantic) noun class and gender systems, but also covers possessive systems (3.3.2.3), morphological systems (3.3.2.4), and phonologically determined noun classes (3.3.2.5).

3.3.1 Noun classification

3.3.1.1. Noun classifiers

Noun classifiers are separate lexemes that appear together with a noun in certain syntactic environments (Dixon 1986: 105; Sands 1995; Wilkins 2000). The (p.75) following example is from Kayardild and includes the classifier (generic) noun *wanku* ‘elasmobranch fish’:

Kayardild (Evans 1995: 244)

(37)	dathin-a	dangka-a	niya	wumburung-kuru	raa-ja
	that-NOM	man-NOM	3SG.NOM	spear-PROP	spear-ACT
	wanku-ya		kulkiji-y		
	elasmobranch-MLOC		shark-MLOC		

‘The man speared a shark with the spear’

Noun classifiers are typically found in Australian languages (such as Kayardild) and were characterized as follows in Dixon (1980: 102–3; but see also e.g. Foley 1986: 78 on noun classification in the languages of New Guinea):¹⁴

Australian languages all have a number of generic terms; a speaker of Yidiny, for instance, will gloss *jarruy* as ‘all the birds’, *narra* as ‘any sort of vine’, and so on. Many languages include the appropriate generic noun in each noun phrase, together with a specific noun. In Yidiny, for example, felicitous sentences can be translated (showing generic nouns in italics): ‘The *bird* crane stood on the *ground* sand’. ‘The *person* woman girl picked up some *stone* slate’, and so on. There are in fact two types of generic noun: (i) classifying specific nouns according to the inherent nature of their referents—*jugi* ‘tree’, *munyimunyi* ‘ant’ and those just exemplified; (ii) classifying specific nouns according to the function or use of their referents—thus Yidiny includes *bona* ‘any drinkable liquid’, *mayi* ‘any edible plant’, *gugu* ‘any purposeful noise’ and *wirra* ‘any moveable object (which is not governed by other generics of type (ii))’.

Dixon (*ibid.*) adds that the use of these classifier (generic) nouns is ‘somewhat reminiscent’ of numeral classifiers, to which we will turn next.

3.3.1.2. Numeral classifiers

In certain languages NPs containing a cardinal numeral may have an extra constituent: the numeral classifier (CLF), which is so called because it typically occurs with a numeral and is deemed to provide some sort of classification of the head noun.

Thai (Allan 1977: 286)

(38) khru·	lâ·j	khon
teacher	three	CLF: person
‘three teachers’		

In the sample the following languages may be characterized as numeral classifier languages: Burmese (Sino-Tibetan), Mandarin Chinese (Sino-Tibetan), Korean (Altaic), Hmong Njua (Austroasiatic), Nivkh (language isolate), Nung (Austroasiatic), and Vietnamese (Austroasiatic).

(p.76) The phrase ‘numeral classifier’ is perhaps somewhat misleading.¹⁵ Although it is generally assumed that the classifier originally occurred only with the numeral (as in e.g. Burmese; Lehman 1990: 105), its occurrence can also spread to other constituents, the demonstrative being among the first. Furthermore the choice of the numeral classifier is often a consequence of the way the speaker conceives of the image of a particular referent in a certain context. Consequently, in different contexts different choices may be made (see section 3.3.1.2.1 below). Or, as Adams and Conklin (1973: 2–3) put it: ‘Shifting of the numeral classifier of a noun shifts the focus from one salient parameter to another [note omitted]’ (cf. also Adams 1989: 3).

Because of their association with a particular constituent, numeral classifier systems are usually distinguished from noun class systems proper (Dixon 1986). Nevertheless the boundary between numeral classifier systems and noun class systems can be rather fuzzy, especially if the two systems are related diachronically (Greenberg 1972, 1978a).

3.3.1.2.1. Variation in classifier choice

Many authors have observed that a noun does not necessarily occur with the same numeral classifier.¹⁶ In Mandarin Chinese, for instance, a goat may be referred to as *yì-zhīyáng* ‘one-animal goat’, *yì.tóu yáng* ‘one-head goat’, *yì-tíao yáng* ‘one-long-thing goat’, or *yì.ge yàng* ‘one-general-classifier goat’ (Erbaugh 1986: 400; see also e.g. T’sou 1976; Li and Thompson 1989: 112). As another illustration consider a well-known example from Burmese (Sino-Tibetan), which has no single classifier for *myi?* ‘river’. According to Becker (1975: 113) ‘the choice depends upon the universe of discourse. One might speak of a river in at least eight contexts’:

(39) <i>myi? tə ya?</i>	‘river one place’ (e.g. destination for a picnic)
<i>myi? tə tan</i>	‘river one line’ (e.g. on a map)
<i>myi? tə hmwa</i>	‘river one section’ (e.g. a fishing area)
<i>myi? tə 'sin</i>	‘river one distant arc’ (e.g. a path to the sea)

myi? tə θwε	'river one connection' (e.g. tying two villages)
myi? tə 'pa	'river one sacred object' (e.g. in mythology)
myi? tə khu'	'river one conceptual unit' (e.g. in a discussion of rivers in general)
myi? tə myi?	'river one river' (the unmarked case)

(p.77) 3.3.1.2.2. Number of classifiers

Languages differ significantly in the number of numeral classifiers that they employ. At one end of the scale there are languages using only one numeral classifier (e.g. Cebuano), and at the other end we find languages that are reported to use 200 or more (e.g. Vietnamese, Burmese). The Cebuano classifier 'is a general classifier that can be applied to almost anything' (Jones 1970: 2). An additional property of the Cebuano classifier is that it is obligatory when the numeral expression follows the head noun, but optional when it precedes (*ibid.* 10); this appears to be rather typical if the numeral may either precede or follow the head noun (cf. e.g. Greenberg 1975: 40 on Ainu; Comrie 1981a: 269 and Gruzdeva 1998: 23–5 on Nivkh). Another peculiarity of the Cebuano classifier phrase is that a so-called linker occurs in between the cardinal numeral and the numeral classifier; a linker also occurs on e.g. adjectives and demonstratives:¹⁷

Cebuano (Jones 1970: 2)

(40) upát-ka buqúk bátaq
four-LIN CLF children
'four children'

Mokilese has only four numeral classifiers, but even with so few of them a noun may occur with more than one classifier:

Mokilese (Dixon 1982: 221)

(41) a. woi pah-men b. woi pah-w
turtle four-CLF turtle four-CLF
'four turtles' 'four turtles'

Most sea creatures can take either *-men* or the general classifier *-w* (general classifiers occur in many numeral classifier languages and are employed instead of the more specific classifiers). It remains unclear what determines the choice of *-men* or *-w* in *woi pahmen* and *woi pahw* 'four turtles'; it is stated that there is no difference in meaning.

Generally languages use many more classifiers, but it is often difficult to determine the precise number (cf. Dixon 1982: 216; Austerlitz 1980: 238). This is at least partly due to the fact that in some languages, such as Burmese, it is possible to create new classifiers (Becker 1975: 113; see also Dixon 1982: 233 on Thai). There are at least two other factors that make it difficult to establish the number of classifiers: (i) the same form may occur both as a noun and as a numeral classifier, and (ii) numeral classifiers may also be used anaphorically.

3.3.1.2.3. Optionality of numeral classifiers

Apart from the fact that classifiers tend not to occur when the numeral modifies a noun denoting some kind of measure or unit (like the equivalents of English ‘inch’ or ‘week’), there are also (p.78) certain contexts in which they are absent with nouns that otherwise do require their presence. This is because in the process of grammaticalization the classifier may take on a wide range of secondary functions which mostly have to do with referentiality in that the classifier is (also) used as a marker of definiteness, topicality, specificity, or genericity (Bisang 1996, 1999; see also section 5.2.2.1.3.2).

In Mandarin Chinese, for example, the numeral classifier never occurs in non-referential or generic NPs; if an NP contains a classifier it must refer to a specific referent (Li and Thompson 1989: 130; Sun 1988). Similarly, in Vietnamese only syntactically referential NPs containing a numeral must contain a classifier (Löbel 1999). Thus the classifier is absent in (42), which is a generic statement, but not in (43):

Vietnamese (Löbel 1999: 297)

(42) Con chó có bốn chân				
CLF dog have four leg				
‘A dog has four legs’ (lit. ‘A dog is four-legged’)				
(43) Con chó tôi cô bốn cái chân ló'n				
CLF dog I have four CLF leg long				
‘My dog has four long legs’				

3.3.1.2.4. Constituents that collocate with numeral classifiers

Numerical classifiers are assumed to have occurred originally with numerals only, but in many numerical classifier languages the classifier also collocates with other NP constituents. For instance, in Mandarin Chinese it also occurs with demonstrative pronouns (Li and Thompson 1989: 130–2; Jones 1970: 11), and in Thai numerical classifiers collocate with numerals (including ordinals), demonstratives, and adjectives (Hundius and Kölver 1983: 172; see also Lehmann 1982a: 255). The Thai numerical classifier is only obligatorily present with numerals; with the other constituents its occurrence seems to be determined by the degree of emphasis placed on the constituent in question (Jones 1970: 7; cf. also Hundius and Kölver 1983: 173 ff.; Bisang 1996, 1999).

3.3.2 Noun classes

Languages with properties associated with prototypical noun class systems (see the introduction to section 3.3 above) do not form a homogeneous group, just like the group of languages using a system of noun classifier (section 3.3.1). They may differ in various respects. For instance, some languages have only two genders/classes, others may have more than a hundred; in some languages gender is associated with number

(e.g. Krongo) and/or size or shape (e.g. Alamblak, Nama), in others this is not so. These and other properties of noun class systems are discussed below.

(p.79) 3.3.2.1. Humanness and animacy

It was mentioned earlier that humanness and animacy, along with the feature Shape (Friedrich 1970), are probably the most important parameters along which spatial entities are classified. This is also reflected in the fact that in some languages only humans or animates trigger distinct forms of expression (see also e.g. section 3.2.5). In the sample at least Abkhaz, Basque, Hixkaryana, Hungarian, and Sarcee belong to these languages (see also section 3.2.4 on numerals in Galela).

In Abkhaz the human distinction plays a role in the selection of the plural marker and in the numeral system in general (see also section 2.2.1). Compare, for example:

Abkhaz (Hewitt 1979: 149, 152, 236–7)

- (44) a. à-jyab b. á-jyab-c° a

ART-girl ART-girl-PL

- (45) a. à-là b. a-la-k° a

ART-dog ART-dog-PL

When marked as genitives or locatives, Basque demonstratives occur with the affix *-ta-* for inanimates and the suffix *-gan-* for animates (Saltarelli 1988: 215–16).

In Sarcee the ±human distinction (as well as number) is reflected in the form of the demonstrative pronoun, but only in the forms that are used to indicate location near the speaker; *díní* ‘this’ (+Human SG), *díná* ‘these’ (+Human PL), *dìyí* ‘this, these’ (–Human SG/PL). Compare:

Sarcee (Cook 1984: 73–4)

- (46) a. *díní ts'iká?* ‘this young woman’

b. *díná ts'ikúwá?* ‘these women’

c. *dìyí tsí* ‘this paint’

Cook (1984: 75) suggests that *díní* is related to the personal pronoun *ídíní* (3SG), and observes that *díná* is identical to the noun meaning ‘person, people’.

Sometimes there is a correlation between humanness/animacy and nominal subcategories (or *Seinsarten*). For instance, in Igbo (Niger-Kordofanian) inanimate nouns seem to be set nouns (i.e. the unmarked noun may be used to refer to a set containing one or more individuals), which can optionally be provided with a collective aspect marker (*ńdi*). Non-personal animate nouns on the other hand seem to be singular object nouns, since the unmarked noun is used to refer to a single entity, unless provided with a plural marker (Welmers 1973: 220).

Hixkaryana does not have a regular number marking system. Instead it has a ‘collective marker’, which is used when the group character of the referent is emphasized.

Collectivity is marked primarily with humans, but also with animals and items that are regarded as an integral part of the culture or environment (p.80) (Derbyshire 1979: 126). If a collective marker is used in the NP, it usually also occurs in the verbal complex.

In Hungarian, numerals optionally occur with the suffix *-an* or *-en* (depending on vowel harmony considerations) when they relate to humans. There is, however, evidence to suggest that this form is an adverb rather than a numeral (Greenberg 1975: 39–40).¹⁸ Finally, in Koasati plurality (if it is that; see section 4.2.1.2) is only marked in the case of humans (Kimball 1991: 447).

3.3.2.2 Exhaustive noun class systems

There are many genetically unrelated languages in which each noun belongs to a certain class or gender, which according to Dixon is one of the prototypical characteristics of a noun class system (cf. Table 3.1). This is the case in at least the following languages in the sample: Alamblik, Babungo, Dutch, Ket, Krongo, Nama Hottentot, Ngalakan, and Oromo. They will be discussed briefly below. Languages that share typically gender-related properties will be taken together.

3.3.2.2.1. Dutch, Ket, Krongo, Ngalakan, and Oromo

All nouns in the three southern Oromo dialects Boraana, Orma, and Waara are either masculine or feminine. Generally nouns that end in a long vowel are masculine, whereas those ending in a short vowel are feminine (Stroomer 1987: 70).¹⁹

Dutch has two genders: a common gender (which comprises the erstwhile masculine and feminine gender) and a neuter gender, which is reflected in the form of the definite article in the singular (*de* for nouns with common gender; *het* for neuter nouns) and in the form of certain adnominal modifiers. Attributive adjectives take an *-e* suffix, except (i) before a singular neuter noun preceded by an indefinite article and some other words (e.g. *elk* ‘each’, *geen* ‘no’, *menig* ‘many a’, *zo’n* ‘such a’), and (ii) before a singular neuter noun preceded by nothing (e.g. *goed bier* ‘good beer’). Compare (*boom* ‘tree’ has common gender and *kind* ‘child’ has neuter gender):

Dutch

- (47) a. een oud-e boom b. de oud-e boom

a big-E tree the big-E tree

‘an old tree’ ‘the old tree’

- c. (de) oud-e bomen

(the) big-E trees

‘(the) old trees’

(p.81)

- (48) a. een klein kind b. het klein-e kind

a small child the small-E child

‘a small child’ ‘the small child’

c. (de) klein-e	kinder en
the	small-E children
'(the) small children'	

Ket also has a system of three noun classes (masculine, feminine, and neuter) which reveals itself primarily in predicate agreement, the forms of the spatial cases (cf. Werner 1994, 1997a; see also section 3.2.5), and the possessive prefixes. The masculine, feminine, and neuter correlate to a large extent with animate male/female and inanimate, but some animal names are always feminine or masculine, irrespective of the sex of the referent (for instance, most fish names are masculine). There is some correlation with size: large wooden referents are masculine, but small ones are neuter (Comrie 1981a: 263).

All Ngakakan nouns are distributed over four classes (Merlan 1983: 235): masculine, feminine, and two inanimate classes—all of which are characterized by a particular prefix. One inanimate class (marked by the *gu-* prefix) includes among other things most body-parts, the majority of the specific tree names, and some general term referring to vegetation. The other class (the *mu-* class) includes many terms for edible and inedible plants. Krongo also has four genders or classes, but this system is discussed in more detail in section 3.3.2.2.4 on the reversal of gender between the singular and the plural of nouns (*polarity*).

Note finally that in some languages there is a correlation between gender and case. For example, in Modern Greek most masculine nouns end in an -s in the nominative singular and have no ending in the genitive-accusative case; conversely, most feminine nouns have a genitive-singular form ending in -s and no ending in the nominative-accusative (Joseph and Philippaki-Warburton 1987: 152; cf. Plank 1979: 624 on a similar phenomenon in Old French).²⁰

3.3.2.2.2. Babungo, Kisi, and Bukiyp

In Babungo, Kisi, and many other African languages, nouns are assigned to different classes on the basis of different patterns of agreement. Thus Babungo has fourteen classes, but the same noun may occur in different classes. This is due to the fact that the class distinctions in Babungo and other Bantu languages take into account number distinctions as well. For example, class 1 contains a certain group of nouns because they all trigger the same kind of agreement pattern when they are used in the singular, and the same group belongs to class 2 because they also trigger the same kind of (p.82) agreement pattern when they are used in the plural; thus pairs of Bantu noun classes can be joined into a gender or class in the sense used here. Probably most nouns belong to two ‘classes’ (the double-class genders), one for the singular and the other for the plural form (Denny and Creider 1986). Compare these examples from Sesotho; numbers indicate noun classes:

Sesotho (Demuth et al. 1986: 456)

(49)	Mq-thq	é-mq-holq	ó-rata	ø-ntjá	□-ntle	ea-haq
1	1 1	1	9	9	9 1	
	person	big	3SG-like	dog	beautiful	of-his/her

'The old man/woman likes his/her beautiful dog'

(50)	ba-thq	bd-ba-hqlq	ba-rata	li-ntjá	ts□-ntlε	tsá-bona
	2	2 2	2	10	10	10 2
	people	big	3PL-like	dogs	beautiful	of-them

'The old people like their beautiful dogs'

In Babungo the most common double-class genders are 1/2, 3/4, 3/13, 5/6, 5/13, 7/8, 9/2, 9/10, and 19/6a. Some nouns only belong to one class (the single-class genders; e.g. what Schaub 1985: 183 calls event nouns such as 'happiness' or 'fog' which belong to class 1 only), whereas others may even belong to three classes (the triple-class genders). For instance, a number of nouns that are used to refer to body-parts occurring in pairs (breasts, cheeks, etc.) belong to classes 5, 6, and 13 in Babungo. Class 5 is used when reference is made to one out of the pair, class 6 is used to refer to the pair, and class 13 is used for more than one pair (Schaub 1985: 182).

Kisi has seven distinct noun classes; roughly, six can be combined into three singular-plural pairings and one contains only liquids and has 'singulars' ('a drop of ...') in one of the singular classes (namely the inanimate class; Childs 1995: 148 ff.).

Bukiyip resembles the African noun classes described above in that nouns are organized in singular-plural pairings. Bukiyip has thirteen of them and, as is also typical the case in the sub-Saharan type system of noun classification in Africa, many noun modifiers must agree in class with the head noun:

Bukiyip (Conrad 1991: 9, 12)

(51) a.	bú-b	b.	bú-bús
	betel nut-CLI.SG		betel nut-CLI.PL
	'betel nut'		'betel nuts'

(52) a.	Iowa-nú	b.	Iowa-nab
	tree snake-CL6.SG		tree snake-CL6.PL
	'tree snake'		'tree snakes'

3.3.2.2.3. Alamblak and Nama Hottentot: gender replacement

In Alamblak (Indo-Pacific) and Nama Hottentot (Khoisan) nouns usually occur with the same (p.83) gender marker (either masculine or feminine), but they can also appear with the opposite gender marker. These languages share some other characteristics as well. In both languages gender is expressed in a phrase final person-number-gender (PNG) affix (which suggests it is rather a property of the NP or its referent than of the noun) and both languages have a so-called indefinite gender.

Alamblak. NPs in Alamblak obligatorily contain a Person-Number-Gender (PNG) marker, an element that shows great similarity with a corresponding member of the set of personal pronouns. As the name suggests, this phrase-final enclitic element indicates number, person, and, in the case of third person singular, also gender:

(53) a. yima-m	b. yima-kë	c. yima-nëm
person-3PL	person-2PL	person-1 PL
'people'	'you people'	'we people'
(lit. 'they people')		

The third person singular has two distinct forms for the masculine and the feminine gender *yima-r* 'man', *yima-t* 'woman'), but some nouns always occur with the same gender marker. These include names of individuals and of natural objects believed to have originated from humans. For instance, *mar-r* 'sun' is masculine, because in local folklore the sun is the son of the moon (*yam-t*).

Most nouns can take either gender and divide into two subgroups: those which are used to refer to entities whose sex is relevant (humans and higher animals), and another group that is used to refer to entities for which sex is not relevant. Nouns of the first subgroup are assigned a gender on the basis of the sex of the referent. Nouns of the second subgroup are marked for gender on the basis of a secondary or extended meaning of the gender marker. That is, the masculine gender marker (-r) can be used to refer to typically tall or long, slender or narrow entities; the feminine gender marker (-t) may be selected to refer to short, squat, or wide entities. This is the unmarked use of the gender system.

However, nouns of either subgroup may occur with the opposite gender. When this happens with nouns of the second subgroup (where sex is irrelevant), this means in the case of inanimates that the size of the referent is atypical. Thus, *kuñ-*'house' is commonly feminine as a consequence of its squat shape, but an unusually long house will be referred to by an NP headed by *kuñr*, ending with the masculine PNG marker. In the case of lower animals the use of the PNG marker means that the sex is highlighted.

Notice that this system is difficult to characterize, as indeed are many of the systems described here. Normally every Alamblak noun has either masculine or feminine gender, which is typical of a grammatical noun class system. If the opposite gender is used, this is on the basis of characteristics of the referent (sex, (p.84) unusual size) as conceived by the speaker, which is rather like the system found in e.g. numeral classifier languages.

Finally, if for some reason the speaker cannot or will not indicate the gender/sex of the referent, the third person plural is employed as an 'indefinite gender marker' (Bruce 1984: 98; cf. also Corbett 1991: 218–19).

Nama Hottentot. Nama NPs are characterized by the fact that they end in a word that is inflected for person, number, and gender (Hagman 1974: 40 f., 84 f.). In this context two basic groups of nouns are distinguished: animate and inanimate nouns. Animate nouns are inflected according to the biological sex of the referent, masculine or feminine: /írí-p 'the male jackal', /íris 'the female jackal', *kxòe-p* 'the male person', *kxòe-s* 'the female person'. Thus, an animate noun is assigned a gender which correlates with the biological sex of the referent. (As a consequence of their meaning certain nouns invariably occur with the same gender marker, such as *’áop* 'the man or husband', *tarás* 'the woman or wife').

An inanimate noun has either masculine or feminine gender, but the rationale behind this twofold division is obscure (Hagman 1974: 44) and any such noun may appear with the opposite gender marker to indicate that there is something unusual about the referent.²¹ It is difficult to provide a common meaning to this kind of gender alternation, because it depends on such diverse factors as the noun in question, the extra-linguistic context, etc. Nevertheless (Hagman 1974: 45):

We may generalize by saying that gender replacement conveys the meaning 'largeness of size with derogation' when largeness is an undesirable characteristic of the referent of the noun stem, it conveys simply 'largeness of size' when largeness is neither desirable nor undesirable, and it may even convey the meaning 'smallness of size' if smallness is undesirable.

There is also an 'indefinite gender', which is not assigned to any noun in the lexicon; any noun can appear with this gender, which replaces the masculine or feminine. Its use indicates that the referent of the NP is unknown, hypothetical, or even non-existent (for more details, see Hagman 1974: 46 f.).²²

(p.85) 3.3.2.2.4. Krongo and Oromo: polarity

In some languages there is an interesting association between gender and number, in that certain nouns have one gender (e.g. feminine) in the singular and another (e.g. masculine) in the plural. This phenomenon is also known as polarity: the reversal of gender between the singular and the plural of nouns (G. Hudson 1976: 252).²³ The sample contains two such languages, namely Krongo (Niger-Kordofanian) and Oromo (Afro-Asiatic). Polarity is also attested in other languages of the Kadugli group of the Kordofanian branch of the Niger-Kordofanian phylum (Tucker and Bryan 1966: 304–5), in languages of the Nilo-Saharan phylum (e.g. Teso, East Sudanic branch; Tucker and Bryan 1966: 467), and seems to be rather common in the Afro-Asiatic family (Bender et al. 1976b: 145), especially in the languages of the eastern group of the Cushitic branch of which Oromo is a member (Serzisko 1982). Although the languages involved belong to three different phyla in Ruhlen's (1987) classification (but cf. Schadeberg 1981), they are all spoken in approximately the same area (Sudan, Ethiopia, Kenya; see also section 9.2.3 on certain other areal features of the languages of Central and East Africa).²⁴

To give an example, Krongo has three genders in the singular. These gender distinctions are not marked on the noun itself, but on the demonstrative, the numeral, the adjective, and the predicate:

Krongo (Reh 1983: 46)

(54)	m-áðlyà	káaw	m-áðéelá	m-ó?ò
	F-come	person	F-nice	F-this
‘this beautiful woman is coming’				

In the case of humans and various domestic animals gender correlates with the sex of the referent. With other referents the gender (masculine, feminine, neuter) is usually fixed, but in the case of non-domestic animals gender may be changed so as to have it correlate with the sex of the referent (Reh 1985: 132–3).²⁵

In the plural one of three things may happen as regards gender marking: (i) the agreement pattern remains the same; (ii) the agreement pattern changes to one that is used with one of the other genders (as used in the singular; hence ‘polarity’); (iii) the agreement pattern that only occurs in the plural is employed (here the term ‘plural gender’ is used; see also Welmers 1973). Which one of these options is selected depends on the noun in question. Not every gender (p.86) combination is attested in the singular/plural opposition. Only the following combinations occur (Reh 1985: 127):

Table 3.2. Gender/number in Krongo

Singular	Plural
Masculine	Masculine
Feminine	Masculine
Feminine	Plural gender
Masculine	Plural gender
Neuter	Masculine
Neuter	Feminine
Neuter	Neuter
Neuter	Plural gender

In some languages one could even classify nouns on the basis of the way they form the plural alone. This has been suggested for a number of Nilo-Saharan languages, in which there is an enormous amount of variation in the formation of plurals (Welmers 1973: 139). However, it remains to be seen whether these plural markers are not actually collective aspect markers (see section 4.2.1).

3.3.2.3 Possessive constructions

In many languages there are similarities in the expression of existential, locative, and possessive constructions (see also Chapter 6). Apparently the same distinctions are more or less relevant in connection with (real or apparent) systems of noun classification. In section 3.2.1 we saw that in some languages nouns are classified on the basis of the positional verb with which they occur in an existential, locative, or possessive construction. In the subsequent sections we will see that in some languages nouns are divided into two classes on the basis of the feature <± Alienable>.²⁶ Alienable possession is typically characterized in terms of ownership. In the case of inalienable possession the ‘possessed entity’ is conceived of as being inherently related to another entity (physically or cognitively). Inalienably possessed items generally obey the following hierarchy: body-parts <kinship relations < part/whole < clothing, tools <other (Seiler 1983; Croft 1988: 172). That is, if in a language tools are regarded as inalienably possessed items, then so are kinship relations and body-parts. Although one could argue that the distinction between alienable and inalienable possession is due to the way referents are conceived by the language user (which would imply that we are dealing with an apparent noun class system), it is generally assumed that the (p.87) noun that is used to refer to an inalienably possessed entity has two arguments (rather than one). In this view (in)alienability is a feature of the noun, which means one can also classify nouns in terms of the number of arguments they take (for ease of reference nouns marked <+Alienable> will be called alienable nouns and nouns

carrying the feature <–Alienable> will be called inalienable nouns; Harrison 1988: 66). Thus, alienable nouns take one argument, whereas inalienable nouns take two (but cf. Mackenzie 1987b). Inalienable nouns are also called ‘relational nouns’, because they typically designate a property of a referent in relation to another referent. For instance, the noun ‘father’ designates a property of a spatial entity, who is of course at the same time the father OF someone (Mackenzie 1983).

3.3.2.3.1. Inalienable possession

It is usually not possible to refer to an inalienably possessed entity without also making reference to some other entity (the possessor). Thus, there is a certain group of nouns which may be used to refer to entities that cannot occur in their own right, i.e. they are two-place, relational nouns. In such cases one cannot refer to just ‘milk’, ‘an arm’, or ‘a mother’, but only to e.g. ‘her milk’, ‘your arm’, or ‘their mother’. The difference between alienable and inalienable nouns may be expressed in various ways. First of all there is of course the fact that inalienable nouns necessarily require reference to a possessing entity, whereas this is not the case for alienable nouns. However, the difference may also manifest itself in other ways. Below are some examples.

Jukun (Niger-Kordofanian) has different sets of possessive pronouns, which correlate with alienable and inalienable nouns; compare (see also e.g. Kiefer 1985: 108 on Hungarian and Haas 1941 on Tunica (Amerind)):

Jukun (Welmers 1973: 218–19)

(55) INALIENABLE	ALIENABLE
zò m̄ my friend	bè m̄ my money
zò ú your (SG) friend	b` bú your (SG) money
zò á his friend	bè bá his money
zò í our friend	bè bí our money
zò ní your (PL) friend	bè bu ní your (PL) money
zò bé their friend	bè bu bé their money

The forms in the second column, the more complex (marked) set, are those that occur with alienable nouns. Cross-linguistically it seems that, in terms of expression, the inalienable construction is normally the unmarked category.²⁷

(p.88) In some languages the realization of the possessor pronoun (free vs. bound form) correlates with alienable and inalienable possession. For instance, in Burushaski possessive *prefixes* are restricted to inalienable nouns (body-parts and kinship relations; section 6.3.1).

In the Papago (Amerind) certain possessive construction always take the suffix *-ga*, whereas others never occur with this suffix.²⁸ ‘These nouns [i.e. without *-ga*] consist of inherently possessed things, such as body-parts and kinship terms, as well as nouns which are said to be inalienably possessed, such as clothing and utensils’ (Zepeda 1983: 79; see also Bahr 1986); compare:

Papago (Zepeda 1983: 79)

(56) Husi we:nag [inalienable]

Joe brother/sister

'Joe's brother/sister'

(57) Huan kawyu-ga [alienable]

John horse-GA

'John's horse'

Ngiti has a small group of nouns that is exclusively found in inalienable possessive constructions, only few of which can be cited in isolation. As a rule the human nouns must be preceded by *ale-* 'person', and the non-human nouns by (*i*)*dhu-* 'thing' (i.e. what belongs to non-humans, including animals):

Ngiti (Kutsch Lojenga 1994: 138)

(58) *alɛ-dɔ*

person-head

'a person's head'

(59) *dhu-dɔ*

animal-head

'the head of an animal'

Inalienable possession is expressed by juxtaposition of the head noun and the possessor NP, whereas the expression of alienable possession requires the appearance of a possessive marker (*bhà* or *tɔ́*):

Ngiti (Kutsch Lojenga 1994: 138)

(60) kamà bha dza

chief POS house

'the chief's house'

(61) kamà-dɔ

chief-head

'the chief's head'

(p.89) As in Jukun there are two sets of possessor pronouns. The set used for inalienable possession follows the head noun, whereas the set used for alienable possession precedes (see also 5.2.1.1):

Ngiti (Kutsch Lojenga 1994: 139)

(62) pbàkà ikyì

1SG.POS cow

'my cow'

(63) afi-du

heart-1SG.POS

'my heart'

In some languages it is possible to ‘alienize’ otherwise inalienable nouns so as to emphasize the absence of a specific or definite possessor. One such language is Saker (also called Bargam; Indo-Pacific), which uses the prefix *ka-* for this purpose:

Saker (Ultan 1978c: 26; from Z'graggen 1965)

(64) ya i-nen

I 1SG-mother

'my mother'

(65) ka-nen

KA-mother

'mother'

In this context it may be appropriate to call attention to a phenomenon that is often referred to as ‘possessor deletion’ (cf. Seiler 1983: 18 f.). Contrary to what one might expect, the (usually inalienably) possessed noun may occur without reference to the possessing entity in certain languages. The following examples are from Haya (Niger-Kordofanian):

Haya (Hyman 1977: 99)

(66) η-k-óogy

émikôno

I-PAST-wash

hands²⁹

lit. 'I washed hands' (= 'I washed my hands')

(67) ? η-k-óogy

émikôno yange

I-PAST-wash hands my

'I washed my (detached) hands'

According to Hyman (*ibid.*) this phenomenon is restricted to certain nouns (e.g. body-parts) and depends on the nature of the verb (notably those involving an experiencer), among other things.

(p.90) Sometimes it is possible to divide inalienable nouns into two subtypes on the basis of formal differences, as in Onondaga (Amerind). In this language inalienable nouns that are used to refer to body-parts ‘that are under voluntary control of the

possessor ... are assigned subjective pronominal prefixes and the locative morpheme' (Woodbury 1975b: 33–4; see also Seiler 1983: 21 f.). Such entities are, for instance, 'head', 'foot', and 'back'. Inalienable nouns that are used to refer to 'uncontrollable body-parts', such as 'intestines', 'blood', and 'hair', take an objective pronominal prefix and lack the locative suffix. Compare (*ibid.*):

Onondaga (Woodbury 1975b: 33–4)

(68) keshúh-ne

my_back-LOC

'my back/on my back'

(69) akyú:œ?

my_intestine: SX

'my intestines'

Finally, there is evidence to suggest that particularly inalienably possessed nouns (such as those designating body-parts) are involved in constructions referred to in connection with so-called possessor ascension (e.g. Fox 1981; Blake 1984). In this view there is a structural relationship between the following (Dutch) sentences:

Dutch

(70) Ik kus-te haar hand

I kiss-SG.PAST 3SG.F.POS hand

'I kissed her hand'

(71) Ik kus-te haar de hand

I kiss-SG.PAST 3SG.F the hand

'I kissed her hand'

In the first sentence the possessor is referred to by means of a possessive pronoun (*haar* 'her') that is part of the object NP. In the second the same referent is now 'promoted' (in terms of the hierarchical strata of Relational Grammar), functioning as a distinct participant in the event referred to.

3.3.2.3.2. Alienable possession: possessive classifiers

In Fijian and other Oceanic languages alienable possessive constructions are characterized by the occurrence of a classificatory element: the possessive classifier (or: relational classifier; Lichtenberk 1983). The following examples are from Mokilese (see (p.91) Seiler 1983: 35–9 on possessive classifiers in Amerindian languages): Mokilese (Harrison 1988: 66)

(72) nimoai pil

CLF[DRINK.1SG] water

'my water (for drinking)'

(73) oai pil

CLF[GENERAL.1SG] water

'my water (for washing)'

Although these examples indicate that variation of the possessive classifier is possible this seems to be the exception rather than the rule. According to Harrison (*ibid.*)

the 'generally accepted analysis' may be rather exaggerated. In my experience with Mokilese at least, paradigmatic classifier contrasts such as those exemplified above are somewhat atypical, and often rather contrived. That is, the classifiers can be *made* to contrast, but there are very few domains for which such contrasts are normal and productive. Moreover, most of those instances in which a classifier is used without a possessed noun are open to an anaphoric analysis.

On the whole the Micronesian languages are reported to have more classifiers than other Oceanic languages. Boumaa Fijian has only three (Dixon 1988: 135 ff.), Mokilese has fourteen, and Trukese is reported to have an open set of possessive classifiers, 'though the number in common use is probably not far from the Micronesian average' (i.e. between fifteen and twenty; Harrison 1988: 66–7).

3.3.2.4. Morphological noun class systems

In some languages morphological subclasses are distinguished. This is the case, for instance, in Onondaga (Amerind), which, according to Woodbury (1975b: 17 ff.), has three morphologically distinct types of nouns. We will ignore one of these, the 'syntactic nouns', because it seems to consist of derived predicates and idiomatic expressions.³⁰

The first type is characterized by the fact that the nouns are uninflected and cannot be incorporated (unless they are nominalized). Some examples are *tsí:ha* 'dog' and *takós* 'cat'.³¹

The so-called morphological nouns constitute the second type. They consist of a noun root with a neuter or animate pronominal prefix and with a noun suffix (p.92) (Sx; see also below). Recall, however, that Onondaga is an Iroquoian language and in at least some of the languages in this family, such as Cayuga and Tuscarora, nouns constitute at best a minor word class. In other words, it may be the case that under a different analysis constructions such as those given in the examples below are treated as clausal structures and that a more appropriate (literal) translation of *o?néhsa?* 'sand' would be 'it sands' (cf. note 30):

Onondaga (Woodbury 1975b)

(74) o?néhsa?

it_sand:SX

'sand'

(75) haksá?a

he_boy:SX

'boy'

Within this group of morphological nouns, two subclasses are distinguished. One subclass consists of nouns that take the neuter prefix; these nouns are used to refer to 'natural' or 'man-made' entities. 'Natural nouns' usually take objective pronominal prefixes (O), and the pronominal prefixes of the subjective set (S) occur with 'man-made nouns'; compare:

Onondaga (Woodbury 1975b)

(76) otsyú?ta?

O.N:fish:SX

'fish'

(77) kahú:wa?

S.N:boat:SX 'boat'

The other subclass within the group of morphological nouns contains nouns that take animate prefixes. This class can also be further divided into kinship and non-kinship nouns; only the latter can be incorporated.

3.3.2.4.1. Noun markers

In the sample Gude is an example of a language that has a group of nouns that must sometimes appear with a semantically empty element. This phenomenon is attested in languages across the globe (such as Niger-Kordofanian, Austric, Afro-Asiatic, Nilo-Saharan, Amerind) and has been described and explained by Greenberg (1977, 1978a, 1981, 1991a; cf. also Childs 1983), mostly in relation to Niger-Congo languages. In these articles Greenberg put forward a theory about the way nouns may have acquired gender/class. According to this theory, four stages can be recognized in the process. Initially (Stage 0) there is a demonstrative pronoun, which may or may not be marked for gender or class. In the following stage (Stage I), the demonstrative has lost its (p.93)

Stage Ø	→	Stage I	→	Stage II	→	Stage III
dem (+ cl)	→	def.art. (+ cl)	→	gen.art. (+ cl)	→	class marker
dem (- cl)	→	def.art. (- cl)	→	gen.art. (- cl)	→	noun marker

Fig. 3.1. Stages in the acquisition of class/gender or noun markers; dem = demonstrative, def.art. = definite article, gen. art. = generic article, ±cl = ±classified (after Greenberg 1978a)

deictic function and has in fact turned into a definite article. In Stage II the article indicates specificity, among other things.³²

We may define Stage II as the stage in which we have an article which includes, along with possibly other uses, both definite and indefinite specific uses. (Greenberg 1978a: 62)

At this stage the articulated form [i.e. the noun-plus-Stage II article] is heavily predominant in text, but there still remain a number of constructions with the non-articulated form so that most nouns may occur in both forms. These non-articulated uses derive from the two ends of the determination spectrum. They may lack the article because they are automatically definite, e.g. vocatives, proper nouns, and possessed—particularly inalienably possessed nouns, such as kinship terms. On the other hand the unarticulated form often survives in various generic uses, typically as negative objects, as predicate noun or adjective, in adverbial, particularly locative, expressions, e.g. ‘on foot’, as incorporated noun objects, and as dependent members of nominal compounds. (Greenberg 1981: 106)

Finally, in Stage III, the element becomes a gender/class marker, but only if the erstwhile attributive demonstrative pronoun was marked for gender or class. In the other case the form will end up as a noun marker, ‘a mere sign of nominality on the large majority of common nouns’ (Greenberg 1978a: 69) (see Fig. 3.1).³³ In the sample Stage II and Stage III (or Stage II/III) articles occur in Gude and outside the sample for instance in Mandinka (Niger-Kordofanian), Mangbetu (Nilo-Saharan), and Fijian (as well as other Oceanic languages; see also section 6.2.1.3 on so-called articles in Abkhaz, Galela, and Hungarian).

Gude. In Gude every noun belongs to one of two morphological subclasses, which the author labels ‘free stems’ and ‘captive stems’ (Hoskison 1983: 22 ff.). Nouns of the captive class must occur with the suffix *-nə* in most contexts, (p.94) whereas nouns of the free class never have this semantically empty element suffixed to their stem. The suffix does not appear before plural, demonstrative, and definiteness suffixes. It may also not appear before a possessive suffix, except in the case of inalienable possession. However, if the inalienably possessed noun follows a preposition and if it is unmodified, it, too, will appear without the suffix.

The relevance of the morphological noun class system for the proper expression of the NP as a whole is also illustrated by the fact that it determines how the noun is pronounced in different contexts and what stem changes occur when suffixes are added.

Hoskison, who also mentions Greenberg’s (1978a) theory discussed in the previous section, could find no obvious correlation with a semantic class. He also states that in some closely related languages (Fali of Vimtim, Fali of Bahuli) nouns are distributed differently over the two subclasses; on the other hand, all nouns of Fali of Bwagira appear to belong to the captive class (Hoskison 1983: 24; cf. also Newman 1990: 18).

Mandinka. In Mandinka (Niger-Kordofanian) any noun can occur in the ‘Stem form’ or in the ‘O-suffix form’ (Rowlands 1959: 37 f.). Rowlands compares the function of the O-suffix with the function of the definite and indefinite articles in English. In isolation Mandinka common nouns—but not proper names—are always given in the suffix form, which is usually characteristic of languages with Stage II articles (Greenberg 1978a: 63). An interesting property of this suffix is that it appears on the following adjective, if there is one, and may co-occur with a demonstrative. Apparently the o- suffix also functions as some kind of focus marker; according to Rowlands (1959: 150–6) it has the same effect as a certain degree of stress in English (see also section 10.2.2.3).

Gambian Mandinka (Rowlands 1959: 65)

(78) nyīng dínding-o

this boy-0

'this boy'

Mangbetu. In Mangbetu (Nilo-Saharan) there is an article-like prefix on nouns, called *déterminateur* [determinator] by Larochette (1958), which also expresses singular or plural number. It is omitted in certain grammatical contexts, e.g. after certain prepositions, with proper names, in the vocative, in inalienable possessive constructions, etc. In sum, it is not used 'lorsque le nom est suffisamment détermine' [when the noun is sufficiently determined] (Larochette 1958: 33; cf. also Tucker and Bryan 1966: 36, 56). Evidently the Mangbetu prefix closely resembles Greenberg's Stage II article.

Noun (phrase) markers (Stage III articles) in Oceanic languages. Greenberg (1977, 1978a, 1981) illustrated his theory about the acquisition of gender or noun markers for the most part with examples from languages spoken on the African and American continent, but it seems that, to a certain extent, a similar process (p.95) can be reconstructed for the development of noun markers in languages from the Pacific, notably in those that belong to the Oceanic branch of the Austric phylum (cf. Greenberg 1978a: 60).

In many Oceanic languages NPs contain an element that is often obligatorily present and which is called an article (Anceaux 1952) or noun phrase marker (NM). The following examples are from modern standard Fijian (Crowley 1985: 136–7).

Fijian (Crowley 1985: 136–7)

(79) na no-gu vale

NM CLF-1SG house

'my house'

(80) na bilo tii

NM cup tea

'cup of tea'

(81) na bilo ni tii

NM cup for tea

'tea-cup'

(82) Au n̩unu wai

1SG drink water

'I drink water'

(83) Au n̩unu-va na wai

1SG drink-TR.3SG NM water

'I am drinking the water'

Notice that the marker does not occur in *Au ?unu wai* 'I drink water' where the noun is said to be incorporated in the transitive verb, and that only the NP as a whole takes the marker. It has been suggested by Crowley (1985: 176–7) that the noun phrase marker was originally 'a marker of a specific or a definite noun phrase (or, something semantically close to either of these functions). This reconstruction ties in reasonably well with Harrison's (1982) reconstruction of [a formally identical element] with a deictic function in western Austronesian languages.'

In several respects the noun phrase marker is rather similar to the equivalent of Greenberg's Stage II article. For instance, in Boumaa Fijian the 'article' can often be translated into English by a definite or an indefinite article, it is absent when the noun is incorporated and it cannot occur in certain syntactic environments (Dixon 1988: 114f.).³⁴

In this connection it is perhaps interesting to note that in some languages, such as Lenakel, the noun phrase marker has become fused with the head noun, so that (p.96) there is 'a two-way noun class system, in which nouns are distinguished by whether or not they carry the prefix *n-* in their syntactically unmarked forms' (Crowley 1985: 167; cf. Greenberg's Stage III article). It is unclear if this morphological noun class system has consequences for the proper expression of the NP or the sentence.

3.3.2.4.2. Declension classes

In languages like Latin nouns are grouped into different declension classes, according to their case endings. Compare for instance the forms in Table 3.3 (the colon after a vowel V indicates a long vowel of quality V). Besides belonging to a certain declension class, a Latin noun is also classified according to its gender: masculine, feminine, and neuter. Although there is considerable overlap between these two systems of noun classification, there are also many exceptions (cf. also Corbett 1991: 34 f. on Russian). For instance, nearly all nouns in declension class I are feminine, but *poeta* 'poet', which has masculine gender, also belongs to declension class I. This can be shown in the form of the adjective, which agrees in gender:

Table 3.3. Some declension classes in Latin

	Decl. I	Decl. II
Nominative	femin-a	hort-us
Genitive	femin-ae	hort-i:
Dative	femin-ae	hort-o:
Accusative	femin-am	hort-um
Ablative	femin-a:	hort-o:

Latin

(84) a. femin-ae bon-ae 'to the good woman'

- b. poet-ae bon-o ‘to the good poet’

In sum, Latin nouns are classified both grammatically, which is reflected in the gender system, and morphologically, which is reflected in the declension classes.

3.3.2.5 Phonological noun class systems

In the previous sections we saw that nouns are classifie

- grammatically (semantically), i.e. based on inherent lexical features of the noun;
- morphologically, in that in certain languages
(p.97)
 - (a) some nouns are formally characterized in certain grammatical environments by some semantically empty element (namely the noun marker), which is supposed to have originated from a demonstrative pronoun (Greenberg 1978a; Crowley 1985);
 - (b) each noun belongs to a certain declension class;
- syntactically, in that certain nouns (i.e. bivalent nouns) commonly require reference to another entity as well.

It was shown that these class or gender systems co-determine the form or order of constituents in a linguistic expression. The present section will show that there are also noun class systems which operate on a phonological basis. Such systems are attested, for instance, in languages of the Torricelli and Sepik families, where in some languages nouns are classified on the basis of their ‘phonological stem’ (Foley 1986: 85).³⁵ For instance, in his description of Arapesh, Fortune (1942: 6–7) proposes several rules to produce the correct plural forms, each of which covers a noun class, e.g.:

1. A noun ending its singular in *b* ends its plural in *bys*.
2. A noun ending its singular in *bɸr* ends its plural in *rɸb* or *ryb*.
3. A noun ending its singular in *g* ends its plural in *gas* or *as*.
4. A noun ending its singular in *ku* ends its plural in one of the following variants: *-meb, rib, ib, guhijer, hijer, ibijer, jeriu, heu, u, omi*.
5. A noun ending its singular in *m* ends its plural in *eip* or in *ip*.

According to Fortune (*ibid.*) ‘these rules cover thirteen classes of nouns, which are fully productive in syntax, each class having its own peculiar form of singular and plural, which pronoun, adjective and numeral must follow in concord’. To some extent there seems to be an overlap between phonological class and gender in that membership of one phonological class implies masculine gender and membership of another phonological class feminine gender.³⁶

Yimas (Sepik-Ramu, Indo-Pacific) has a dozen noun classes; four are semantic (male humans, female humans, higher animals, plants, and certain plant products) while the other eight are phonologically determined (Foley 1986: 86). Every noun must be marked for class and number and every other NP constituent obligatorily agrees with the head noun in number and class:

Yimas (Foley 1986: 85–6)

(85) ama-na-ŋki-i	triŋk-i	mamaki-ŋk-i
1SG-POS-CL6-PL	tooth: CL6-PL	bad-CL6-PL

(p.98)

k-ramnaw-t	k-ia-k
CL6-three-PL	CL6-PL-this
‘these three bad teeth of mine’	

The concord affixes in Yimas are divided into two types, basically the adjectival set and the verbal set. Both types for class 6 are illustrated above. The adjectival set are suffixes, and generally of the same phonological form as the final consonant or consonant cluster of the stem for the vast bulk of inanimate nouns whose class membership is determined phonologically. In the above example *triŋk* ‘tooth’ is a class 6 noun, signalled by the stem ending in *-ŋk*. The possessive and adjective modifiers must agree in class with *triqk*, and so they are suffixed with the adjectival suffix for class 6 *-ŋk* and the plural morpheme *-i*. The other two modifiers occur with a different morpheme, the prefix *k-*. This is the verbal agreement marker for class 6. Yimas verbs agree in class and number with their central nominal participants, subject, object, and indirect object. The verbal agreement markers are always prefixes, corresponding to the final consonant and sometimes to the preceding vowel of the noun stem.

Another way in which nouns are classified on a phonological basis is attested in certain Bantu languages (e.g. Lomongo, Lingala, and Kinyarwanda; see also Sasse 1981 on the Cushitic languages), where nouns can be classified according to their different tone patterns (Daeleman 1983). Thus, the short CVCV-stems can be divided into four tone groups (L = low tone, H = high tone): LL, LH, HL, and HH. For instance (high toneme is indicated by an acute; glosses and translation of these examples from the KiNtándu dialect of Kikongo are not provided in Daeleman’s article),

Kikongo (Daeleman 1983: 133)

(86) ma-lafu	(LL)
ma-tutí	(LH)

In some such languages the tone pattern of each tone group changes according to the function of the noun in the sentence. This phenomenon is referred to as tone case. That is, if the noun occurs in a subject or object NP its tone pattern will be different from the one used when the noun functions as the sentence predicate; there are four tone cases in all. Cf. (˘ indicates a short rising toneme):

Kikongo (Daeleman 1983: 133)

(87) ma-lafu	(LL; subject/object)
ma-łafu	(HL; predicate)

(88) ma-tuti (LH; subject/object)

ma-túti (HL; predicate)

(p.99) 3.4. Conclusion

In this chapter I have tried to show (necessarily in a rather arbitrary fashion) how real and apparent noun class distinctions can be relevant for the proper expression of linguistic structures. First I gave examples of the effect of class systems on constituents *outside* the domain of the NP (section 3.2): sentence predicates and the relators (adposition, case affix). Then I demonstrated how such systems can influence the form and order of constituents *inside* the domain of the NP and following Dixon (1986) I distinguished between noun classifiers and noun classes. After considering some general properties of *classifier systems* (which includes both noun classifiers and numeral classifiers), I focused on the more traditionally recognized *noun class systems* (including genders), involving such semantic features as <±Human> (including <Masculine>, <Feminine>, <Neuter>) or <±Animate>. We also saw that in certain languages the distinction between alienable and inalienable possession determines the expression of the NP (as regards form/order), which may be regarded as a classification that is based on the number of *arguments* that a noun takes. There are also languages in which nouns can be classified on a *morphological* basis, in that e.g. a particular group of nouns usually occurs with an affix (which may not appear under certain conditions; for instance, when the NP contains a demonstrative or possessive pronoun). And finally I mentioned some languages in which nouns can be classified *phonologically*. In such cases the phonological form of the noun proves to be relevant in connection with e.g. plural marking (Arapesh) and the tone pattern of the noun in the sentence (Kikongo).

Up to this point I have mainly focused on properties of the noun, the central constituent of the NP. The next three chapters are concerned with properties of modifiers of the noun in the NP such as determiners, numerals, and adjectives. Needless to say, the determination of the fundamental properties of these modifier categories is a prerequisite for an understanding of the overall structure of the NP. Moreover, the properties of noun modifiers are of prime relevance in connection with the ordering principles to be discussed in Chapters 8–10. To get a clear idea of linearization within the NP, we need to establish for the languages in the sample first of all whether they actually employ such categories as articles, numerals, adjectives, etc. It is also important to know whether these constituents are part of the integral NP; if they are appositional constructions or distinct phrases at the level of the clause, their position in the NP is not (necessarily) determined by the ordering principles considered in Chapters 8–10. For the same reason we need to know whether or not a particular noun modifier is expressed as a free element; the position of constituents relative to the head noun is also taken into account so as to be able to test certain hypotheses that can be formulated on the basis of the Principle of Head Proximity (Chapter 9).

Notes:

- (1) A semantic classification of nouns in terms of *Seinsarten* was given in Chapter 2; recall also that this study is mostly restricted to first order nouns that are used in connection with real-world individuals (section 1.5). Some important contributions in

the area of gender and other forms of noun classification are Fodor (1959); Goral (1978); Greenberg (1978a); Seiler and Lehmann (1982); Craig (1986b); Corbett (1989); Senft (2000b); see Mithun (1999: 95–103) on gender in the languages of native North America. The reader is particularly referred to Aikhenvald's monograph on classifiers (Aikhenvald 2000) and Corbett's book on gender (Corbett 1991). However, due to the fact that Corbett regards agreement as the sole criterion, certain manifestations of noun classification are ignored, such as differences in ordering properties between alienable and inalienable nouns. Furthermore, he does not distinguish between properties of nouns and properties of referents of NPs (or, rather, the ontological counterpart of the referent in the real world).

(2) Not surprisingly, a division into grammatical genders is absent in languages without a distinct class of nouns, such as Samoan (V/N/A), Quechua, Turkish, and Hurrian (all N/A), and Cayuga (V(-N)); see also section 2.2.4.

(3) In some languages the form of the sentence depends on the kind of entity referred to. For instance, in Spanish the copula *estar* is used with first order (or spatial) entities, as in *la mesa está en la sola 14* [the table COP in the room 14] ‘the table is in room 14’), whereas *ser* is used with second order (or temporal) entities, as in *la réunion es en la sola 14* [the meeting COP in the room 14] ‘the meeting is in room 14’ (Hengeveld 1986, 1990b).

(4) On inalienable possession see also Ameka (1996); Evans (1996); and other contributions in the collection of articles in Chappell and McGregor (1996).

(5) Cf. Boas (1911a: 38–9); Lang (1975); de Vries (1989: 194); Piau (1981–2); Wurm (1982: 81); Foley (1986: 88–90); Merlan et al. (1997). It has been argued that there is a diachronic relation between positional verbs and classifiers; see e.g. Rankin (1977); Watkins (1976b); Barron and Serzisko (1982).

(6) The relationship between existential and locative (as well as possessive) constructions has been investigated in various studies; see section 6.1 for references.

(7) In Dutch the progressive construction without a positional verb but with the verb (*te*) *zijn* ‘(to) be’ is also available: *Hij was de krant aan het lezen* [he was the newspaper at the reading] ‘he was reading the newspaper’.

(8) Nouns denoting body-parts may or may not be incorporated in Onondaga but this always involves a difference in meaning (Woodbury 1975b: 49). For another example of noun classification based on incorporation, see Allen and Gardiner (1977). Weir (1990: 324) presents a classification of Nadëb nouns in terms of their possessability (unpossessable nouns cannot incorporate).

(9) On classificatory verbs in other Na-Dene languages, see e.g. Hoijer (1945) on the Apachean languages; Carter (1976) on Chipewyan; and Landar (1964) on Navajo (cf. also Davidson et al. 1963; Krauss 1968). See Haas (1948); Denny (1978, 1979); Barron (1980); and Blankenship (1997), for some examples from languages of the Amerind phylum.

(10) See Watkins (1976b: 4) on the verbal origin of classificatory articles in Ponca (Amerind) and Rankin (1977) on the development from positional verb to classificatory element in Siouan languages. See also Moser (1977: 20–1) on Seri (Amerind): ‘There is

a phonological and semantical relationship between the definite articles and certain [copular/positional] verbs.'

(11) See also e.g. Craig (1986a) on Jacalteco and Payne (1987) on Yagua. Grinevald (2000: 77–9) suggests that the systems of nominal classification can be associated with different operators in the layered model of the noun phrase (Chapter 7, see also section 11.2; note that genitive or possessive classifiers are discussed in section 3.3.2.3.2): 'the various classifier types are different operators which correspond to different modes of individuation manifested by their different semantics: noun classifiers and their predominantly material/essence semantics are operators of quality, numeral classifiers with their characteristic physical (shape, texture, size) semantics are operators of quality, while genitive classifiers with their functional semantics are operators of locality.'

(12) Cf. Lehmann (1982a: 255) on the occurrence of the numeral classifier outside the numeral classifier phrase: 'The use of the classifier within the same NP which contains the classified noun is diachronically secondary. The classifier phrase enters into a sort of appositive relationship with the classified noun, the former functioning as the syntactic head of the construction [note omitted]. In the course of proceeding grammaticalization, the classifier phrase becomes a modifier of the classified noun. At the same time, the classifier phrase spreads to other nominal modifiers.'

(13) Recall that classificatory elements may be a reflection of some semantic feature of the nominal predicate or of some conceived property of the referent of the NP. Notice that these are also the two factors that may be responsible for iconicity in syntax: (i) isomorphism, the tendency for there to be a one-to-one correspondence between form and meaning, (ii) motivation, the reflection in linguistic structure of some aspect of the structure of reality (Haiman 1980: 515–16, 1985b; Kirsner 1985: 250).

(14) On nominal classification in Australian languages in general, see Harvey and Reid (1997).

(15) It is therefore better to speak of 'sortal classifier' (Lyons 1977: 463; Dik 1989: 123), as opposed to 'mensural classifier', which is used in connection with mass nouns (see also Chapter 2).

(16) On variation in classifier choice, see for example Adams (1986: 241–6); Burling (1965); Benton (1968); Adams and Conklin (1973); Becker (1975); Denny (1976, 1986); Clark (1976); Allan (1977); Goral (1978); Sellner (1982); Dixon (1982, 1986); Erbaugh (1986); Lehman (1990). See also Blankenship (1997: 95) on variation in the choice of classificatory verb (section 3.2.2): 'The LIVING class need not be rigidly adhered to ... Depending on which quality the speaker wishes to emphasize, a cat can be LIVING or FLEXIBLE, and a tree can be LIVING or LONG.'

(17) 'Linkers' are also called 'ligatures' (Foley 1980).

(18) According to Edith Moravcsik (personal communication) the best *prima facie* evidence to show that the Hungarian numerals with *-an/-en* are adverbs is that the same affix is also a de-adjectival adverbializer, compare: *kedves-en* [nice-ly] 'nicely' and *hat-an* [six-ly] 'six of them'.

(19) In Hausa (which belongs to the Chadic branch of the Afro-Asiatic phylum) nearly all common nouns end in a long vowel. Greenberg (1977: 103) argues that these long vowels are vestiges of former articles (more on this in section 3.3.2.4.1; cf. also Newman 1991: 60).

(20) This can also be seen as an instance of polarity (section 3.3.2.2.4), as denoted by Serzisko (1982: 196): ‘Polarität ist dadurch charakterisiert, daß bei zwei Oppositionen dasjenige Element, das den markierten Pol der einen Opposition kennzeichnet, auch verwendet wird, um den markierten Pol der anderen Opposition anzudeuten’ [Polarity is characterized by the fact that the same element that marks a member in one opposition is also used to mark a member in another opposition].

(21) Often gender may have other meanings (especially in the case of gender switch). For instance, in Tigrinya (Afro-Asiatic) feminine gender does not only indicate female sex, but has also diminutive, pejorative, or affectionate significance (Welmers 1973: 231). See Heine (1982: 198 f.) on the derivative potential of ‘free genders’, which also occur in the Bantu class system (notably in relation with the so-called minor classes (Place, Augmentative, and Diminutive), and which contain almost every noun (cf. also Greenberg 1978a: 79; Corbett 1991: 159–60).

(22) Greenberg (1978a: 79) discusses the same phenomenon in Nama and gives some more examples from other languages (Chinook, Khasi; cf. also Rabel-Heymann 1977). According to him, it shows how a neuter, common gender may arise from an indefinite article. He also quotes Meinhof (1909: 48) who wrote: ‘The -/of the genus commune was originally an indefinite article which had nothing to do with gender and it is still used this way at the present time. It can be added to every substantive whether of masculine or feminine gender to signify an instance of the appropriate class.’ See also e.g. Dimmendaal (1983: 218) on Turkana (Nilo-Saharan).

(23) See notes 20 and 24. The term was coined by Meinhof (1912); see also e.g. Corbett and Hayward (1987), Hetzron (1967: 184ff., 1978a).

(24) Here I restrict myself to polarity in the NP proper, but it also occurs in predicate agreement in certain northern Oromo dialects (Serzisko 1982: 188 f.; Corbett 1991: 195–8).

(25) ‘Some Nouns appear to be assigned to a Gender on the basis of analogy, according to the initial consonant of the Prefix or Stem (the term phonetic gender is used to describe this phenomenon)’ (Tucker and Bryan 1966: 304–5; cf. also Reh 1983: 134 ff.).

(26) On inalienability see the volume edited by Chappell and McGregor (1996); see also Ljung (1974) for the morphological process of adjective derivation in relation with alienably and inalienably possessed nouns.

(27) Cf. Ultan (1978b: 26; see also section 6.3): ‘In short, a morphologically derived form [of the possessive pronoun] for a less intimate possessor implies its derivation from the more intimate possessor form (or base); and a free, less intimate possessor may imply [the existence of] a bound more intimate possessor, but not the converse.’

(28) See also Campbell (1985: 45–6) on Pipil.

(29) More precisely, the Haya prefix *k(a)-* marks ‘before yesterday past tense’.

(30) But cf. Sasse (1988, 1993), who argues that a distinct category of nouns is lacking in Cayuga and other languages of the Iroquoian family (section 1.5.1). In his view there is only a category of verbs, which consists of different subclasses which, in varying degrees, deviate from a group of prototypical verbs, the least prototypical subclass being what he calls the ‘nominoids’.

(31) Sasse (1988: 11–12) writes that many Cayuga nouns of this type are either loans or ideophones.

(32) NPs with specific or generic reference are not treated in any detail in this study (but cf. section 7.7). See e.g. Moreton (1999) and von Heusinger (1999) for recent discussions on specificity and Dahl (1975); Gerstner-Link (1995 and forthcoming); and Behrens (2000) on genericity.

(33) See also Harris (1985: 75 f.) on the Kartvelian languages (Caucasian), where articles are assumed to have developed into case affixes.

(34) Cf. also Crowley (1985: 137); see van den Berg (1989: 98–104) for an account of the (p)article *o* in Muna (Austric).

(35) See section 3.3.2.2.1 on Oromo. Cf. also e.g. Hayward and Corbett (1988: 264) on the correlation between phonological properties and gender in Afar, another Afro-Asiatic language, and Heine (1982: 199 f.) on phonological gender in African languages in general.

(36) For more details see Fortune (1942: 11 f.). On phonological systems in general, see Corbett (1991: 51 f.).

4 Qualifying Modifiers in the Noun Phrase

4.1. Introduction

In Chapter 11 wrote that most noun modifiers can be characterized in terms of the three notions Quality, Quantity, and Locality. This chapter is concerned with elements that relate to the notion Quality in that they specify inherent features of the referent. Cross-linguistically there are two categories that pertain to such features in the NP (besides the head noun, of course). One concerns certain dimensional (spatial) features of the property that is designated by the noun; this is the new grammatical category of *nominal aspect markers* (quality operators) and will be discussed in detail in section 4.2.

The other category involves *lexical* elements that typically specify more or less inherent properties of the referent (quality satellites): *adjectives*. If a language does not have a distinct category of adjectives, it will usually employ qualifying NPs (headed by an abstract noun) or relative clauses (headed by a stative or descriptive verb) for this purpose, as in the English paraphrases ‘the man *with richness*’ or ‘the man *who is rich*’. In section 4.3 I will however mostly focus on specialized quality satellites, namely adjectives (as in ‘*the rich man*’). I will also show that the existence of a distinct class of adjectives in a language is closely connected to the kind of noun type that the language employs. More specifically, I will argue that an open class of adjectives is only found in languages where nouns designate spatially bounded properties (N_{+Shape}).

4.2. Grammatical expressions of the notion Quality in the noun phrase

We saw in Chapter 2 that set nouns are generally characterized by the fact that they are transnumeral and that they usually do not appear in the plural form when modified by a cardinal numeral, as in the Caucasian language Lezgi (Haspelmath 1993: 232): ‘When the cardinal numbers are used attributively, the counted noun is always singular, e.g. *c'ud niik* [ten bird] “ten birds” (**c'ud niik'er*).’

In this section we will see that in many languages the *non-numerated* set noun may or must receive what is often called a plural marker, but I will argue that this element does not express plural number but rather indicates that the speaker refers to a particular kind of set, namely a *collective* set of entities. In some (p.101) languages we even find that the set noun may occur with an element which indicates that the noun designates a property of a singleton set, i.e. a set with only one member. It is important to make clear at the outset that these elements, which I call *nominal aspect markers* and which comprise both *singulative* and *collective* aspect markers, do not express number distinctions but specify what kind of (set) entity is being referred to: a singleton set or a collective set.

4.2.1. Aspect

We can define the notion ‘aspect’ as the way in which a property or relation designated by a predicate is represented in some dimension. Depending on the type of predicate

involved, two kinds of aspect can be distinguished: verbal and nominal aspect. Verbal aspect is concerned with representations in the temporal dimension, and nominal aspect with representations in the spatial dimension. The study of verbal aspect has a long tradition, but nominal aspect has only been introduced recently, at least in the sense in which it is used here.¹ I have already mentioned the difference between covert and overt ways of marking aspectual distinctions (cf. section 2.6 on *Aktionsart* vs. verbal aspect marking and *Seinsart* vs. nominal aspect marking). In this chapter we will only be concerned with overt, inflectional expressions of nominal aspect.²

Let us briefly return to the difference between number marking in Dutch (which typically employs singular object nouns for reference to discrete physical objects) and so-called number marking in Oromo (which uses set nouns). In Dutch the plural marker is obligatory in that it must be used whenever reference is made to more than one individual, both with and without the presence of an adnominal numeral in the NP. In other words, an NP without a plural marker (or some other quantitative specification) refers to a single object (as in 1a) and an NP with a plural marker always refers to multiple objects (as in 1b).

Dutch

-
- (1) a. de/een tafel [the/a table] = one table
b. (de) tafel-s [(the) table-PL] = multiple tables

(p.102) Another characteristic of Dutch (first order) nouns is that the plural marker also appears when the noun is modified by a numeral with a value of ‘two’ or higher:

Dutch

-
- (2) a. (de) twee tafel-s [(the) two table-PL] ‘(the) two tables’
b. *(de) twee tafel [(the) two table]

Since these facts seem to indicate that the unmarked Dutch noun only designates a property of a single object I have called such nouns *singular object nouns*.

In Oromo, on the other hand, the so-called number marker is optional, and must be absent when the noun is modified by a numeral (Stroomer 1987: 76):³

In general, nouns with plural suffixes refer to a counted or countable group of items, whereas the possible plural meaning of nouns unspecified for plural is more general and vague. If a noun is counted by means of a numeral, then there is no plural suffix.

Recall that Oromo nouns are transnumeral in that the unmarked form may be used to refer to one or more entities (see 2.2.1.1).

Oromo (Stroomer 1987: 76–7)

-
- (3) a. farda ‘horse/horses’ vs. fardoollee ‘horses’
b. saree ‘dog/dogs’ vs. sareellee ‘dogs’

Because nouns such as *farda* ‘horse/horses’ and *saree* ‘dog/dogs’ designate a property of one or more individuals and because a set may consist of any number of individuals (including ‘one’), I have called these nouns *set nouns* (section 2.2). Some nouns may also occur with a *singulative suffix*, so there are actually two ways to disambiguate the transnumeral character of nouns in Oromo (Stroomer 1987: 83, 87; BOW = the three Oromo dialects Boraana, Orma, and Waata):

BOW nouns denoting animate beings, in particular ethnonyms, can take the singulative suffixes *-ca* (masculine), and *-tii* (feminine); these suffixes are preceded by the epithetic vowel *i*; *t* is sometimes inserted between the noun root and the singulative suffix. In BOW ethnonyms these suffixes are productive.

In BOW these [singulative] suffixes basically have the meaning of indicating an individual out of a group.⁴

(p.103) Oromo (Stroomer 1987: 84–5)

- | | | | |
|-------------|---------------|---------------|---------------|
| (4) a. nama | ‘man/men’ | vs. namica | ‘a/the man’ |
| b. nad’eeni | ‘woman/women’ | vs. nad’ittii | ‘a/the woman’ |

Now, if both in Dutch and in Oromo we are dealing with a number marker, why do these markers behave so differently? The answer I have proposed (see e.g. Rijkhoff 1991, 2000) is that the Oromo affixes are not number markers at all but grammatical elements indicating that the noun designates a property of a set which consists of one individual (singleton set) or signalling that the set consists of multiple individual entities which together form a collective (collective set). This hypothesis is supported by the fact that in the grammars of languages with set nouns it is often explicitly stated that the so-called plural marker basically has a collective meaning.⁵ Since strictly speaking these elements do not indicate number but rather specify the way the nominal property is represented in the spatial dimension (i.e. they relate to inherent or qualitative properties of the referent) I have called these so-called number markers on set nouns *singulative* and *collective aspect markers*, or more generally *nominal aspect markers*. These elements will be discussed in more detail in section 4.2.1.2, where I will also deal with another phenomenon that is related to set nouns, namely regular instances of so-called ‘number discord’ between argument(s) and predicate.

4.2.1.1. Verbal aspect

So as to be able to appreciate the differences and similarities between verbal and nominal aspect, let us first briefly consider the essence of verbal aspect, already touched upon in Chapter 2. Properties and relations in the temporal dimension, which are designated by verbal predicates (‘sit’, ‘walk’, ‘read’, etc.), can be characterized in terms of two typically temporal features: *Beginning* and *Ending*. This gives us four ways of representing temporal properties and relations, i.e. four basic verbal aspectual distinctions (cf. Table 2.3, which is repeated here for convenience as Table 4.1). Although at least within some of these basic categories finer subdivisions can be made (such as ±progressive in the imperfective and ±momentaneous in the perfective; cf. e.g. Comrie 1976), we may say that generally any property or relation designated by a dynamic verb can be represented in any of these four ways.⁶ When the aspectual distinctions are part of the (p.104)

Table 4.1. Basic aspectual categories of the verb

Time	-Beginning	+Beginning
-Ending	imperfective	ingressive
+Ending	egressive	perfective

lexical meaning of the verb they are studied in the context of *Aktionsart* and when they are overtly expressed by inflectional morphology they are called *verbal aspects*. The following examples of inflectional verbal aspect marking are from Mokilese (Micronesian). The first sentence, with the verb in the imperfective, characterizes the situation as an open-ended event (the chase has not stopped) whereas the sentence with the verb in the perfective describes the situation as a bounded event, i.e. the chase has come to an end (Chung and Timberlake 1985):

Mokilese (Chung and Timberlake 1985: 237)

- | | | |
|--------------------------------|----------------|----------|
| (5) Ngoah kauj-ki | ih | awahioaw |
| I | chase:IMPF-DUR | him hour |
| 'I chased him for an hour' | | |
| (6) Ngoah kauj-kih-di | ih | awahioaw |
| I | chase-DUR-PERF | him hour |
| 'I chased him down in an hour' | | |

Thus, the time adverb has a different sense in these sentences: 'With an imperfective the time expression measures the duration of an open event, while with a perfective it specifies the duration of a closed event' (ibid.; see also the Russian examples in 2.5).

4.2.1.2. Nominal aspect

We saw above that a set noun designates a property of one or more entities and that the nominal aspect marker further specifies what kind of set is involved: a singleton set or a collective set.⁷ Thus the function of the nominal aspect marker is clearly different from the function of a number marker. The (compulsory) number marker on a singular object noun or on a collective noun indicates that we are dealing with multiple singular objects ('chairs') or multiple collectives ('families'); without the number marker the noun designates a property of one (p.105) object ('chair') or one collective ('family'). As I said before, the nominal aspect marker on a set noun (which is usually optional) specifies the kind of set we are dealing with: a singleton set or a collective set. Just as the speaker indicates with a verbal aspect marker how a verbal property or relation is represented in the temporal dimension as regards the features *Beginning* and *Ending*, so he can specify with a nominal aspect marker how a nominal property is represented in space in terms of the features *Shape* and *Homogeneity* (section 2.4.3, Fig. 2.1).

If we accept that the so-called number marker on a set noun is actually a nominal aspect marker, we can explain the differences noted earlier. The fact that the collective aspect marker on a non-numerated set noun is in many languages optional or even altogether absent is not difficult to explain when we realize that the property designated by a set noun also applies to multiple objects without this marker. And the fact that the

collective aspect marker (if used at all) is normally absent when the set noun is modified by a numeral can be explained if we realize that the numeral does not specify the number of sets (as it specifies the number of individual objects or collectives in the case of an individual object noun or a collective noun), but rather the size of the (single) set.

On the assumption that the so-called number marker on a non-numerated set noun is in fact a nominal aspect marker, we can also explain systematic instances of ‘number discord’ between the NP and the verb, a phenomenon that seems to be restricted to NPs headed by transnumeral nouns, in particular set nouns (Rijkhoff 1993; cf. also Moravcsik 1978b: 343–50). This is due to the fact that the pronominal element in the verbal complex may agree with the set (singular verb agreement) or with the individual(s) in the set (singular or plural verb agreement). For instance, in the case of Georgian, Lango, and Oromo ‘number’ agreement is always with the set:

Georgian (Harris 1981: 22)

(7)	sami	knuti	goravs
	three	kitten	roll:3SG
‘three kittens are rolling’			

Lango (Noonan 1992: 168)

(8)	gúlú	àdE ^K	òtòò
	pot	three	3SG:die:PERF
‘three pots broke’			

Oromo (Stroomer 1987: 107)

(9)	gaala	lamaani	sookoo	d'ak'-e
	camel	two	market	go-3SG.M.PAST
‘Two camels went to the market’				

However, ambiguity arises when a numeral is absent and the agreement marker in the verbal complex is transnumeral as well (in that it does not distinguish (p.106) between 3SG and 3PL). This is the case in e.g. Galela with non-human subjects:

Galela (van Baarda 1908: 60)

(10)	o	namo	i	tèmo
	ART	bird	3	whistle
‘the bird whistles’ or ‘the birds whistle’				

Table 4.2 lists a number of semantic and morpho-syntactic features of nouns in languages that in Chapter 2 were identified as employing set nouns (see section 2.2.4 on languages with ‘flexible nouns’, i.e. Samoan (V/N/A), Quechua, Turkish, and Hurrian (all N/A)):

- *noun class*: which nouns belong to the category of set nouns (if known);
- *aspect marker*, whether (non-quantified) set nouns appear with nominal aspect markers (sing = singulative aspect; coll = collective aspect);
- *opt/obi*: whether nominal aspect markers are optional or obligatory;
- *px/sx; -N/N-*: whether they appear as affixes (px = prefix, sx = suffix) or as free elements (before or after the noun, i.e. -N or N- resp.);
- ‘*number*’ *discord*: whether there is ‘number’ discord between argument and verb; here I have only considered the clearest cases: explicit statements in the source about number discord and/or examples of NPs containing a numerated noun that trigger singular marking in the verb.⁸

I have already demonstrated, in section 2.2.2, that all languages in Table 4.2 have a small or large group of nouns (*set nouns*) that do not appear with a so-called number, or rather a *nominal aspect marker*, when modified by a numeral with which they are in a direct construction (as in quasi-English *three house* or *house three* = ‘three houses’). Here I will concentrate on the properties listed in Table 4.2. Much of the evidence will be presented in the form of quotations so as to permit the reader to assess personally the original words of the various authors. Recall however that in my view all so-called number markers on set nouns are most probably instances of nominal aspect marking.

In Abkhaz the plural suffix is usually compulsory but may be absent on animate, non-human nouns if a plural marker appears in the verb. This is especially the case with transitive subject NPs.

Abkhaz (Hewitt 1979: 150)

(11) a-gəgš°èg sə-r-bè-yt
ART-beast me- <u>they</u> -see-FIN
‘The beasts saw me’

(p.107)

Table 4.2. Languages with set nouns

Noun class	Aspect marker	Opt/obi.	Px/sx; -N/N-	‘Number’ discord
<u>Type 3: V-N-A</u>				
Abkhaz	+ anim./- human	coll	opt	sx yes
Basque (NP = INDEF)		coll	opt	sx yes
Berbice Dutch		coll	obi	sx
		sing? and coll	opt	sx (see below)

Burushaski (see below)					
Chukchi					
(case=ABS)		sing and coll	opt	sx	yes
(other cases)		Ø			
Georgian	+ animate	coll	?	sx	yes
	- animate	coll	?	sx	yes
Guaraní		coll	opt	N-	
Hittite	Neuter	coll	opt	sx	yes
Hungarian		coll	obi	sx	yes
Ika		coll	opt	N-	
Kayardild		coll	opt	sx	yes
Ket (see below)	- human/anim.	coll?			
Hmong Njua		coll	opt	-N	
Nasioi(NP = INDEF)	- human	sing and coll	opt?	sx	
Ngalakan	+ human	coll	opt	sx	yes
	- human	Ø			yes
Ngiti	- human				
Oromo	+ human	sing and coll	opt	sx	yes
	- human	coll	opt	sx	yes
Sumerian	+ animate - human	coll Ø	opt	sx	yes
Wambon					
<u>Type 3/4: V-N(-A)</u>					
Gude		coll	opt	sx	
Pipil? (insufficient data)					
Sarcee (see below)	+ human - human	sing? and coll	opt	sx/N-	
Tamil	- human	coll	opt	sx	
<u>Type 4: V-N</u>					
Galela		coll	opt	-N	(see above)
Hixkaryana	+ human - human	coll Ø	opt	N-	(see below)
Koasati	+ human - human	sing/coll (see 4.2.1.3)	opt	sx	
Korean		coll	opt	sx	
Lango		coll			yes

Mandarin Chinese	+ human	coll	opt	sx	
Nivkh:					
if NP = SP	+ animate	coll	obi	sx	yes
else		coll	opt	sx	yes
Nunggubuyu	(see below)				
Tsou	+ animate	coll	opt	px	
<u>Type unknown</u>					
Nahali					

(p.108) In Basque the number marker (which is optional with a numeral) is only obligatory in a definite NP. This is related to the fact that formal expression of number is intrinsically connected with the expression of definiteness and case: these three notions are all expressed in a single portmanteau suffix. Note furthermore that ‘indefinite nominals are sometimes treated as singular and sometimes as plural by the verb agreement system (illustrating the non-distinctive nature of indefinite nominals with respect to the number parameter)’ (Saltarelli 1988: 200).

Basque (Saltarelli 1988: 200)

(12) gizon batzu etorr-i d-i-ra

man some come-PERF 3SG.ABS-PRES-AUX1

‘Some men have come’⁹

Given the fact that in contemporary Basque number discord between NP and predicate is considered archaic (Saltarelli 1988: 201), it may be the case that Basque nouns are currently in transition, changing from set nouns to singular object nouns.

In Berbice Dutch Creole the appearance of the plural suffix *-apu* depends on pragmatic factors in that it only occurs in referential NPs (Kouwenberg 1991: 167; personal communication).

I have already mentioned that the situation in Burushaski is not entirely clear. Nevertheless there appear to be ‘many singular nouns ... which are used with the power of plurals, e.g. they take the verb in the plural’; presumably this involves the same nouns that are ‘used in the singular and plural with singular form’ (Lorimer 1935–8: i. 44). This group includes, for instance, the names of most fruits and some animals. According to Tiffou and Pesot (1989: 17), however, the plural marker is obligatory with original Burushaski words, but optional with most loans. On the other hand, there also seems to be a singulative aspect marker, i.e. ‘a certain element which is probably to be identified with the equivalent of “one” that has the force of a singular suffix when used with a noun or pronoun, of which the form is not definitely plural, but which is habitually, or at least frequently, used with plural force’ (Lorimer 1935–8: i. 48). Cf. also Tiffou and Pesot (1989: 17) on the function and meaning of this singulative suffix:

Le bourouchaski dispose également d'une marque de singulatif *-en*, *-an*. Cette marque peut jouer un rôle analogue à celui de l'article indéfini du français pour isoler un élément d'un groupe ou pour faire une première mention. Le nom ainsi

marqué peut être précédé du numéral *hen*, *han* [see my remark above—JR]; ex: *han gírkisan bim* ‘il y avait un rat’.

(p.109)

[Burushaski has a singulative marker *-en*, *-an* [the form depends on the noun class—JR]. This marker may play a role that is similar to that of the indefinite article in French to isolate an element of a group or to mark a first mentioning. A noun that is marked in this fashion may be preceded by the numeral *hen*, *han* [see my remark above—JR]; e.g. *han gírkisan bim* ‘there was a rat’.]

Georgian commonly uses the suffix *-eb* to express plural number. Harris (1981: 21–2) writes, however, that ‘for many speakers, only animate nominals trigger Number Agreement’ and that ‘nominals that occur with a quantifier are in the singular in Georgian and also fail to trigger Number Agreement’ (cf. also Fähnrich 1986: 50–1).¹⁰

Georgian (Harris 1981: 21–2)

(13) ҡnuṭebi goraven

kitten:PL:NOM rolls:3PL

‘the kittens are rolling’

(14) sami ҡnuṭi goravs

three kitten:NOM roll:3SG

‘three kittens are rolling’

(15) burtebi goravs

ball:PL:NOM roll:3SG

‘the balls are rolling’

In Chukchi number distinctions are only expressed with nouns in the absolute (i.e. zero) case form; for most nouns the number distinction is neutralized with the other cases (Comrie 1981a: 246–8; cf. also Bogoras 1922: 694; Kämpfe and Volodin 1995: 82).

Chukchi (Comrie 1981a: 248)

(16) tum y-e kupre-n na-ntəvat-y?an

friend-INS net-ABS.SG NA-put-3PL.3SG

‘the friends put the net’

In the anti-passive (ANPA) variant the verb *ntəvat* ‘put’ is detransitivized by the prefix *ine-*, ‘friend(s)’ appears in the absolute (*tum y-ət*) and the patient NP *kupre* ‘net’, which is usually omitted, in the instrumental or locative case (*kupre-te*).

Chukchi (Comrie 1981a: 248)

(17) tum	ə-ɣt	kupre-te	ena-ntovat-r?at
		friend-ABS.PL	net-INS ANPA-put-3PL
‘the friends put the net’			

(p.110) Galela is another language in which the bare noun can be used to head an NP that refers to one or more individuals (van Baarda 1891: 12–13; see also van Baarda 1908: 32–3):

Gebruikt men het zelfstandig-naamwoord zonder nadere aanwijzing van getal, dan kan men zoowel het enkelvoud als het meervoud van het benoemde voorwerp bedoelen.

[If one uses the noun without further specification as regards number, then one may imply both the singular and the plural of the designated entity.]

Optionally the noun may be preceded by the particle *bi*, as when the speaker wants to emphasize that more than one individual is involved. However, this particle may never occur with a cardinal numeral:¹¹

Galela (van Baarda 1891: 13)

(18) o	bi	gotta
	ART	PL
tree		
‘trees’		

As is typical for Chadic languages, Gude has several plural suffixes which do not occur with cardinal numerals greater than one (Hoskison 1983: 34, 51). Although it is not made explicit in Hoskison’s grammar whether a plural marker may be absent without a modifying numeral, it has been stated elsewhere that none of the Chadic languages has obligatory plural marking (Frajzyngier 1997: 195–6).

Hixkaryana is reported not to have a regular singular-plural opposition (Derbyshire 1979: 126, cf. also pp. 83–4).¹²

It does, however, have a way of marking ‘collective’ when the focus is on a group ... Collective is marked primarily only with ‘human’ class nouns; this class includes, in addition to human beings, animals and items regarded as an integral part of the culture and environment of the people ...¹³

With nonhuman nouns the number-marking system is obligatorily absent. With human nouns it is optional in the sense that the speaker can decide whether he wishes to focus on the group or not. In both cases there is often not any clearly disambiguating factor present.

It is not clear if *komo* ‘collective’ is also obligatorily absent if the head noun is in construction with a numeral. According to Derbyshire (1979: 154) there are only (p.111) three basic numerals (*towenyxa* ‘one’, *asako* ‘two’, and *osorwawo* ‘three’), which rather function as adverbs. Also, Portuguese numerals have come to be employed recently, and are regarded as nouns. In order to function as

adverbials these numerals are followed by *me*, which Derbyshire calls a denominalizer.¹⁴

Apparently neuter nouns in Hittite could occur with or without the plural ending (Luraghi, personal communication; Friedrich 1974: 117; Drohla 1949). Furthermore only plural NPs headed by a neuter noun occurred with singular verbs (Friedrich 1974: 118; Garrett 1990: 290).

In Hungarian the plural marker is obligatory without the numeral, but ‘after numerated nouns ... only singular verb forms [may be used] in (present-day) Hungarian’ (Moravcsik 1978b: 344).

Hungarian (Moravcsik 1997: 315)

(19) ez a két alma

this the two apple

‘these two apples’

Ika nouns ‘are not marked as singular or plural, but the qualifier *džina* (plural) conveys the idea of “more than one” without specifying any absolute or relative quantity’ (Frank 1990: 29–30).

Ika (Frank 1990: 29)

(20) nΛ-gunam̩ džina

1-worker PL

‘my workers’

In Kayardild number marking is optional as well and only occurs for reasons of emphasis (Evans 1995: 183):¹⁵

Kayardild (Evans 1995: 235–6)

(21) kiyarrng-ka yarbud-a ngarnal

two-NOM meat-NOM white_cockatoo: NOM

(Number) (Generic) (Specific)

‘two white cockatoos’

(p.112)

(22) dathin-a kiyarrg-ka jungarra nal-da banga-a

that-NOM two-NOM big:NOM head-NOM turtle-NOM

‘those two big turtle heads’

Ket provides an interesting case in the present context, for which I have no ready explanation. On the one hand, there is a small group of animate nouns (e.g. *ty·t* ‘mosquito(es)’, *i·s* ‘fish(es)’, and *bΛ?n* ‘duck(s)’) which never appear in the plural and which trigger singular and plural agreement in the verb (in bold print): cf. *i·s’ s’es’ka duraraq* [V:SG] ‘the fish lives in the river’, and *i·s’ s’es’ka duri n* [V:PL] ‘the fishes live in the river’. On the other hand, NPs headed by inanimate nouns, which are marked for number, normally occur with the verb in the singular (Werner 1997a: 189):

Ket (Werner, personal communication)

(23) a.	<i>tiʔs’ ta·vöt</i>	b.	<i>tΛʔ-η ta·vöt</i>
	stone lie:3SG		stone-PL lie:3SG
	‘the stone lies’		‘the stones lie’

Thus, it seems that two different features of set nouns are distributed over two (largely) non-overlapping groups of nouns.¹⁶

The situation in Koasati is described as follows (Kimball 1991: 446–7; recall that Koasati has no adnominal numerals; section 2.2):

As a general rule, plurality of a noun is not marked on the noun itself. Rather, plurality of a noun subject or object is marked on or within the verb on which the noun depends, as in the following example:¹⁷

Koasati (Kimball 1991: 446; LVP = locative verbal prefix)

(24)	<i>okipófka-k o:w-á:y</i>
	whale-S LVP-go about (SG/DU)
‘A whale is swimming about’	

(25)	<i>okipófka-k o:-yomáhl</i>
	whale-S LVP-go_about (PL)
‘There are some whales swimming about’	

If a noun refers to human beings, it may optionally take the suffix *-ha* to indicate plurality. ... In addition, nouns with the suffix *-ha* have been noted to take only the nominative, accusative and autonomous cases ... Animal names can occur with the suffix *-ha* when they refer to people of the clan of the same name, such as *nitá* ‘bear; bears; Bear-Clan person,’ plural *nitahá* ‘Bear-Clan people’.

Notice, incidentally, that the humanness/animacy hierarchy plays an important role in number and nominal aspect marking (cf. Smith-Stark 1974; cf. also (p.113) Comrie 1981b: chapter 9). For instance, in Abkhaz only human nouns are marked for number, whereas nouns of the non-human class seem to be coded for set aspect. And in Hixkaryana all nouns seem to be coded for set aspect, but only nouns of the human class may occur with *komo* ‘collective’ (see above). In the sample the humanness/animacy hierarchy plays a role in connection with aspect marking in at least the following languages: Abkhaz, Hittite, Hixkaryana, Koa-sati, Lango, Nasioi, Ngalakan, Ngiti, Nunggubuyu, Sarcee, Sumerian, and Tamil.

Most Lango nouns referring to body-parts, implements, fruits, trees, and vegetables always appear in the same form, whether they are modified by a numeral or not.

Lango (Noonan 1992: 167)

(26) a. gúlú ácēl	b. gúlú áryō
pot	one
'one pot'	'two pots'

Furthermore subject NPs take the verb in the singular: Lango (Noonan 1992: 168)

(27) a. gúlú òtōō	b. gúlú àdēk òtōō
pot	3SG:die:PERF
'the pot(s) broke'	'three pots broke'

Although Lango also has many human and animate nouns with a distinctive plural form (there is no regular plural formation; Noonan 1992: 166–7), I have still entered Lango as a language with set nouns for two reasons. Even when a noun has a distinctive plural it is not always used and, secondly, it takes third person singular concord in the verb (the plural verb forms are only used with zero subjects and with the pronoun *gín* ‘they’ (Noonan 1992: 167). Therefore I assume that these distinctive ‘plural’ forms are perhaps better analyzed as lexicalized collective forms.

Lango (Noonan 1992: 169)

(28) àwòbe dúcú òdòk	pí	rwòt
boys	all	3SG:go_back:PERF because_of king
‘all the boys went back because of the kins’		

In Nasioi only non-human nouns display properties of set nouns. Furthermore, since the expression of definiteness necessarily involves the expression of number and gender, the set character only reveals itself in indefinite NPs (as in Basque, but also in e.g. Somali, which is outside the sample; cf. Kirk 1905: 97), for it is then that the referent of the NP headed by the (set) noun may involve one or more individuals (Rausch 1912: 113–15).¹⁸

(p.114) Nasioi (Rausch 1912: 114)

(29) a. páva	b. páva-nàva	c. páva-ni
house	house-DEF.SG	house-DEF.PL
‘a house, houses’	‘the house’	‘the houses’

In Ngalakan all nouns are transnumeral but only kin terms may occur with a plural suffix. Also, ‘referentially specific plural (human or animate) nouns are cross-referenced in the verb by plural pronominals. Thus, **nugu-bigur** may demand

interpretation as *the men*, *the man*, *people* ..., or *a man*, depending on verbal cross-reference and other factors' (Merlan 1983: 39, 53, 56). Inanimate nouns as well as animate lower-order nouns on the other hand are always treated as singular 'unless explicitly marked as dual or plural in some particular context' (*ibid.* 89).

In Ngiti all non-human nouns 'can be interpreted as singular or plural, according to the semantics of the context' (Kutsch Lojenga 1994: 133). There is also no singular/plural opposition for non-human nouns in Nunggubuyu; it is not clear to what extent number marking on human nouns is compulsory (Heath 1984: 195–9).¹⁹ The Sumerian plural suffix *-ene* only occurred with animate nouns, but was mutually exclusive with numerals (Thomsen 1984: 48–9; cf. also pp. 59–60):

In indicating number animate and inanimate nouns are treated differently:... An inanimate noun can denote both singular and plural, or better collective, just like the English word 'sheep', *é* is both 'house' and 'houses' or rather 'complex of houses', *gud* is 'ox' as well as 'oxen'/'herd of oxen'. Inanimate nouns have thus no plural forms, but they can be reduplicated and thus denote a totality: *kur-kur* 'all the foreign lands'. As regards animate nouns, the single stem may probably also denote collective ... but in general plural is expressed by the suffix */-ene/...* Reduplication in the sense of totality occurs also with animate nouns.

In Nivkh number marking is only obligatory in the case of a specific NP headed by an animate noun, but it never appears in the context of numerals and quantifiers (Mattissen and Drossard 1998: 14, 59). According to the same source, there is number discord between argument and predicate in that 'plural forms co-occur with singular forms' (*ibid.*; see also Panfilov 1962–5: i. 92, ii. 112 as well as Gruzdeva 1998: 17).

Wambon nouns are probably all transnumeral and 'number' marking is generally absent. The only exceptions are *kap* 'man, person', which can be reduplicated, and kinship terms, which may receive 'the plural suffix *-ngguye*' (de Vries 1989: 35).

(p.115) Sarcee nouns 'do not normally inflect for number, but a small subclass of nouns, primarily those referring to kinship relations, may occur in the combined plural and collective form that is suffixed by *-ká* or *yìná*' (Cook 1984: 1984).

In Tamil both animate and inanimate nouns of the 'non-rational' class are only optionally marked for number.

Tsou nouns are not marked for number either, but may be provided with the collective marker *ma-* which is prefixed to the reduplicated form of the noun. The latter is mutually exclusive with a numeral (Szakos 1994: 68, 162).²⁰ With certain human nouns plurality is indicated by lengthening of the vowel in the first syllable. It is not quite clear if these nouns can also be regarded as true set nouns, since the numeral may occur with and without the plural form, cf. (AP = attributive particle):

Tsou (Szakos 1994: 70)

(30) a. tueu ci mamespirji	b. tueu ci ma-a-mespiñi
three AP woman	three AP women
'three women'	'three women'

Nahali has a plural suffix (*-ta*), but this is mostly omitted, as in this example (Kuiper 1962: 27):

Nahali (Kuiper 1962: 27)

(31) *īr lānā*

two son

'two sons'

Let us now turn to languages in the sample that use numeral classifiers. Although it has been established that number marking is at most optional in languages that employ numeral classifiers (Sanches and Slobin 1973), all classifier language in my sample are deemed to have some form of so-called plural marking. In the case of Mandarin Chinese and Hmong Njua, however, we are probably dealing with nominal aspect markers, whereas in the other cases these so-called plural markers are better categorized as quantifiers like English 'some', 'many' etc. (section 5.1).

The following examples show that Hmong Njua nouns are transnumeral and that numerals require a sortal classifier when they modify a noun:

Hmong Njua (Harriehausen 1990: 117)

(32) *Kuv yuav tsev*

1SG buy house

'I buy a house/(some) houses'

(p.116) Hmong Njua (Harriehausen 1990: 105)

(33) *Kuv pum Peb tug tlev luv*

1SG see four CLF dog big

'I see four big dogs'

Note, however, that classifiers are also used to express singularity and plurality (or rather collectivity; see below):

Hmong Njua (Harriehausen 1990: 117)

(34) a. *kuv yuav lub tsev* b. *kuv yuav cov tsev*

1SG buy CLF house 1SG buy PL house

'I buy the house'

'I buy (the) houses'

Compare also:

(35) a. *kuv lob hoob* b. *kuv cov hoob*

1SG CLF room	1SG PL room
'my room'	'my rooms'

Harriehausen does not explicitly state that the CLF + N construction in (34a) has singular reference (as indicated by the translations she gives in her book), but according to the Miao Language Team (1972: 16) this is indeed the case (cf. also Bisang 1993: 26; Wang 1972: 125): 'With the exception of the Hsiang hsi dialect, classifiers may occur as single modifiers of the noun to indicate definite singularity.'²¹

Harriehausen writes that plurality is expressed by *cov* (1990: 115, 117):

Cov ist eine 'neutrale' Mengenbezeichnung, die alle Kategoriewörter ersetzen kann (daher die Bezeichnung 'neutral'), wenn die Nominalphrase eine Plural-Bedeutung implizieren soll. Somit hat cov die Bedeutung 'viele, mehrere, mehr als ein'... Ohne weitere Modification kann die Pluralbedeutung eines Nomens nur durch cov ausgedrückt werden.

[Cov is a neutral quantifier, which can replace all classifiers (hence the label 'neutral') if the noun phrase is to imply plurality. Thus cov means 'many, several, more than one'... Without further modification the plurality of a noun can only be expressed by cov.]

Cov is actually a group classifier (section 2.3.1.3) which has a collective meaning and which derives from *cɔy*⁵¹ 'bunches or clusters of fruit' (Ratliff 1991; superscript 51 refers to tone marks).²²

(p.117) In Chapter 2 it was tacitly assumed that the employment of a numeral or sortal classifier implies that the language in question has sort nouns (with the features: -Shape and -Homogeneity). Here we see that Hmong Njua classifiers have widened their functional range and now serve as nominal aspect markers, which implies, of course, that Hmong Njua has set nouns rather than sort nouns. Note, incidentally, that the collective marker is mutually exclusive with a numeral, which is also typical for set nouns (section 2.2.2):

Hmong Njua (Harriehausen 1990: 117)

(36) kuv pum peb tug/*cov tlev luv
1SG see four CLF/*COV dog big
'I see four big dogs'

On the basis of the data presented above we may conclude that Hmong Njua is like other classifier languages in that its nouns are transnumeral, but the fact that in this language classifiers are also used to mark both singularity and collectivity is a strong indication that Hmong Njua employs set nouns [+Shape] rather than sort nouns [-Shape], the kind of noun that is normally attested in classifier languages. Singularity is expressed by adding one of the sortal classifiers, collectivity by using the group classifier *cov*:

Hmong Njua

(37) a. N _{Set}	b. CLF.SG N _{Set}	c. CLF.COLL N _{Set}
tsev	lub	tsev cov tsev
'house(s)'	'the house'	'(the) houses'

The other classifier language in the sample that can be regarded as having a nominal aspect marker is Mandarin Chinese, which is discussed in more detail in section 5.2.1.3 (see also Rijkhoff 2000).

4.2.1.3. Size and nominal aspect marking

In a number of languages with set nouns there is a relationship between nominal aspect marking on the one hand, and the size of the referent on the other. I have no ready explanation for this relationship but one could hypothesize that it might (p.118) have something to do with the fact that size implies a spatial boundary (+ Shape), one of the defining features of set nouns.²³ This relationship has been observed in the following languages (notice that not all languages are in the sample): Asmat, Cuna (Greenberg 1991a: 310), Koasati, Krongo (Reh 1985: 158), Ngiyambaa (Wellegga) Oromo,²⁴ and Sarcee (Cook 1984: 105).

For instance, in Ngiyambaa (Australian) nominal roots ‘are intrinsically neutral as to number’ (Donaldson 1980: 99), but one of the ways to specify what Donaldson calls number is by means of a suffix, which is said to indicate both size and number (*ibid.*).

If a noun is to be marked as singular with such a suffix, it must also be marked for size, large or small, and vice versa. If small, the suffix employed will show whether this is because of immaturity or not.

Stroomer (1987: 87), referring to Gragg (1976: 180–2), writes that the Oromo ‘singulative affix’ *-ittii*, which I regard as a singulative aspect marker (see above), also has a diminutive meaning in Wellegga Oromo when it is attached to a masculine noun:

Oromo (Stroomer 1987: 87)

- | | |
|--------------|-------------------|
| (38) a. mana | 'house/houses' |
| b. manittii | 'the small house' |

Consider also the following examples from Asmat, which has two suffixes, *-nakáp* and *-nakás*. The former is reported to indicate both smallness and singular number, the latter smallness and plural number:

Asmat (Voorhoeve 1965: 128–30; see also Drabbe 1959: 55)

- | | |
|-------------|----------------|
| (39) a. cem | 'house/houses' |
| b. cémnakap | 'small house' |

We have already seen that human nouns in Koasati may optionally occur with the suffix *-ha* ‘plural’. Kimball (1991: 460) writes that ‘for pluralization of non-human animate

diminutives, the independent word *lawísta* “small [PL]” is used. Thus the plural of *nitasí* “bear cub” would be *nítá lawísta* “bear cubs”, literally, “little bears”.²⁵

(p.119) In his article on plural words Dryer (198%) probably also includes instances of what I would regard as nominal aspect markers. For example, in some languages the so-called plural word cannot co-occur with a numeral. Interestingly, Dryer found that in a number of cases his category of plural words also includes words meaning ‘small’ or ‘big’ (see also some of the quantifying elements I listed earlier in the context of so-called plural markers in Burmese and Nung). It is tempting to relate smallness to singleton sets (individuals) and bigness to sets containing more than one individual (a collective), but e.g. in Asmat smallness also goes together with ‘plurality’ (or rather collectivity).²⁶

4.2.1.4 The origin of nominal aspect markers

There is evidence to suggest that collective aspect markers have developed from at least three different sources. One has already been mentioned in the previous section: words meaning ‘small’ and ‘big’. In languages with numeral classifiers it is the group classifier that may assume the function of a collective aspect marker (section 4.2.1.2) and finally there are languages such as Bambara in which it is the third person plural pronoun that optionally serves as a collective aspect marker (Brauner 1974: 26–7). Recall, however, that this third person plural in Bambara attaches to the NP as a whole rather than to the noun.²⁷ Note incidentally that in some of the languages mentioned in Dryer’s study on plural words (Dryer 1989b: 875–6) the category of ‘plural words’ also arose from pronouns (e.g. Yoruba, Chamorro, and Ngarinyin).²⁸

As for singulative aspect markers, Greenberg (1981: 109) has suggested that in certain Nilo-Saharan languages an erstwhile non-classifying Stage III article (or noun marker; see also Greenberg 1978a, 1991, and Chapter 6) ‘has received an apparently secondary interpretation as singular, or perhaps better singulative in relation to a collective. This was already noted by Lukas [1928] in regard to(p.120) Kanuri *kam* “person”; *am* “people”.²⁹ A similar development may have occurred in Anejom (or Aneityum, an Oceanic language of the Austric phylum), where an erstwhile demonstrative or definite article became a noun (phrase)/specificity marker, and now indicates ‘singularity’, or rather singulative aspect (Crowley 1985: 165–6, 178–9; see also Westermann 1912: 26, 46, on the origin of the singular affix in Shilluk).

Anejom (Crowley 1985: 166)

(40) a. n-atam ^w an b. ilpu-atam ^w an	
N-man	ILPU-man
‘man’	‘men’

4.2.1.5. Some morpho-syntactic properties of nominal aspect markers

Nominal aspect markers, i.e. singulative and collective aspect markers, may be bound elements (suffixes, prefixes) or free elements, although in languages where such markers are expressed as—more or less—free elements (such as Galela, Guaraní, Hixkaryana, Sarcee; see section 10.2) there is usually some controversy with respect to their morpho-syntactic status.³⁰

For instance, Galela *bi* ('collective aspect' in my interpretation) is called a particle, which, if present, must always immediately precede the noun. Guaraní *hikwā̄;i* and *kwéra* are lumped into the category of qualifiers, which is basically a wastebasket category (Gregores and Suárez 1967: 144). Hixkaryana *komo* 'collective' is also called a particle (Derbyshire 1979: 83):

Particles are like clitics in that they are phonologically bound to their head word, but they are considered a distinct word class on the grounds that (i) they are mobile compared with bound affixes, and (ii) at their boundaries they are not subject to the same morphophono-logical processes that apply at morpheme boundaries within words.

A similar situation obtains in Sarcee (Cook 1984: 65); witness:

[Sarcee] nouns do not normally inflect for number, but a small subclass of nouns, primarily those referring to kinship relations, may occur in the combined plural and collective form that is suffixed by *-ká* or *yìná* [endnote: *yìná* never occurs alone as a noun, but it is not an affix either ...].

(p.121) Not only is it sometimes difficult to determine whether or not nominal aspect markers are affixes, clitics, or distinct words, but they also seem to represent a category on the boundary between derivational and inflectional morphology, a boundary which is difficult to draw anyway (cf. Booij and van Marie 1996).³¹ It is clear, though, that of all inflectional processes, those that concern nominal (or verbal) aspect come closest to changing the meaning of the property designated by the predicate, a feature that is commonly associated with derivational morphology. However, it will be remembered that strictly speaking aspect does not concern the property or relation as such, but rather the way this is represented in some dimension (time, space), in terms of certain well-established parameters (Beginning and Ending; Homogeneity and Shape).

4.2.1.6. Conclusion

In section 4.2.1 have argued that in many languages so-called number markers are better analyzed as nominal aspect markers. Number markers and nominal aspect markers share the property that they are only used with nouns that can be in a direct construction with a numeral. The difference is, however, that number markers only occur with nouns that designate a property of a single object (or a collective), whereas nominal aspect markers occur with transnumeral nouns, which designate a property of one or more objects, i.e. a set. There are two kinds of nominal aspect markers: singulative aspect markers, which indicate that the set consists of one individual, and collective aspect markers, which indicate that the set consists of more individuals. Furthermore, number markers are compulsory, both with and without an adnominal modifier. Nominal aspect markers, on the other hand, are not compulsory without a numeral and are normally absent with a numeral in the NP. The most important semantic difference is that a number marker specifies the number of individual objects (or collectives), whereas the nominal aspect marker specifies a *qualitative* property of the referent, namely the kind of set the speaker is referring to: a singleton set or a collective set.

4.3. Lexical expressions of the notion Quality in the noun phrase: adjectives

We now take up the issue of *lexical* noun modifiers that pertain to more or less inherent properties of the entity as defined by the head noun. Such properties typically involve categories such as size, weight, color, age, and value and are typically expressed by adjectives. Adjectives, however, are not attested in every language. Just as languages may do without, for example, articles (Chapter 6), numeral classifiers (Chapter 5), or nominal aspect markers (or *qualifying operators*; section 4.2), there are also languages that have no attributive adjectives (*qualifying satellites*). So as to be able to decide whether or not a language has (p.122)

Type 1	V/N/A		
Type 2	V	N/A	
Type 3	V	N	A
Type 4	V	N	—
Type 5	V	—	—

Fig. 4.1. Parts-of-speech systems (based on Hengeveld 1992b)

adjectives, one needs a definition that has sufficient cross-linguistic applicability and is at the same time specific enough to single out the same word class (if it exists) in each individual language. We saw in Chapter 1 (section 1.5.3) that such a definition is provided by Hengeveld (1992b), who characterizes adjectives as predicates which, without further measures being taken, can be used to modify a noun. I will repeat Hengeveld's classification of parts-of-speech systems here for convenience (Fig. 4.1). Recall that in certain languages some or all lexemes can occur in more than one function (types 1 and 2, i.e. languages with so-called 'flexible' lexemes), whereas in other languages there is a strict division of labor as regards the function of the various predicates (types 3, 4, and 5, i.e. languages with rigid or specialized lexemes). Here I am mostly interested in languages of types 1, 2, and 3, i.e. languages that have a class of lexemes (whether flexible or rigid) that fit Hengeveld's definition. Sometimes, however, straightforward classification according to Hengeveld's criteria was not quite possible. For example, in Chukchi 'bare' adjectives (if they are that; see below) only appear when incorporated, whereas in a language such as Tsou it is difficult to recognize adjectives as a distinct category, since verbs do not always require special measures either when used attributively. These and other more or less problematic cases will be discussed separately in section 4.3.3.

4.3.1. Languages with adjectives

Samoan, the only language of type 1 (V/N/A) in the sample, and Quechua, one of the three type-2 languages (V-N/A), have already been discussed in Chapter 1; the other type-2 languages in the sample are Hurrian and Turkish. As to the latter, it has been noted in many grammars that there are no good reasons to make a sharp distinction between nouns and adjectives in Turkish (Lewis 1967: 53 f.) and (p.123) other Turkic languages (Menges 1968: 107; cf. also Poppe 1964: 9 on Altaic languages in general).

Evidence to support the classification of the extinct language Hurrian as a type-2 language is perhaps not so straightforward and is mostly based on statements in Speiser (1941). For example, he writes that the Hurrian noun ‘assumes a number of functions which other languages distribute frequently among separate parts of speech’ and that lexemes that would traditionally be called nouns and adjectives ‘take the same affixes to express attributive relations’ (Speiser 1941: 72). Finally, since he claims that any subdivision between nouns and adjectives ‘is not indicated by any morphological feature’, I take it that Hurrian nouns and adjectives constitute a single flexible word class.³²

Next we turn to the languages that appear to have a distinct class of adjectives according to Hengeveld’s definition. The following languages all have at least two dozen or so lexemes that can immediately be used as noun modifiers: Abkhaz, Alamblak, Basque, Berbice Dutch Creole, Bukiyp, Burushaski, Dutch, Georgian, Guarani, Hittite, Hmong Njua, Hungarian, Ika, Kayardild, Ket, Nama Hottentot, Nasioi, Ngalakan, Ngiti, Sumerian, and Wambon. Apart from superficial adaptations that are due to agreement phenomena, no special measures (participialization, relativization, etc.) are necessary in these languages to let the lexemes in question serve as a noun modifier.

In Abkhaz adjectives follow the noun (probably in a polysynthetic construction; see section 10.2.2.1), except those expressing nationality:

Abkhaz (Hewitt 1979: 56)

(41) à-jyab ḥàrak'

ART-girl tall

‘the tall girl’

Alamblak has approximately forty-five adjectives which appear to cluster in semantic groupings as regards positional preferences and possibilities. Those relating to the notion ‘value’ tend to occur after the head noun, whereas for adjectives relating to the notions ‘physical property’, ‘age’, and ‘human propensity’ the position before the noun is preferred. Adjectives of dimension do not seem to have a preferred position and color adjectives are exclusively attested before the noun (Bruce 1984: 78). Recall that NPs end with an enclitical ‘person-number-gender marker’ (section 3.3.2.2.3); compare

Alamblak (Bruce 1984: 90)

(42) a. ind bro fèh-r b. ind fèh bro-r
DEM big pig-3SG.M DEM Pig big-3SG.M

'the big pig'

'the BIG pig'

(p.124) In Basque it seems only a handful of adjectives may precede the head noun; the rest must appear in postnominal position (Saltarelli 1988: 75).

Basque (Saltarelli 1988: 247)

(43) mutiko biurri hori

child naughty that

'that naughty child'

Basque (Saltarelli 1988: 75)

(44) a. amerikar hiri-a b. hiri amerikar-ra

American city-SG.ABS

city American-SG.ABS

'American city'

'American city'

Adjectives in Berbice Dutch must appear before the noun:

Berbice Dutch Creole (Kouwenberg 1991: 111)

(45) én bam jujέrma

one pretty young_woman

'a pretty young woman'

Adjectives in Bukiyp agree in noun class and occur on either side of the noun.

Bukiyp (Conrad 1991: 37)

(46) a. yopu-kwi élmato^k b. yopu-nali élman

good-CL4.SG woman

good-CL7.SG man

'good woman'

'good man'

The following examples are from Burushaski, Dutch, Georgian, Guaraní, and Hungarian. Recall that Dutch attributive adjectives take an *-e* suffix, except in indefinite noun phrases headed by a *singular neuter* noun (see Holmberg and Rijkhoff 1997 for instances of this phenomenon in other Germanic languages).

Burushaski (Lorimer 1935–8: i. 107; see also Berger 1998: 183)³³

(47) šu.a hirʌn

good man

'a good man'

Dutch

- (48) a. een klein huis Neuter_gender b. een klein-e auto Common_gender
a small house a small-E car
'a small house' 'a small car'

(p.125) Guaraní (Gregores and Suárez 1967: 148)

- (49) kisé piahú
knife new
'a new knife'

Hungarian (Moravcsik 1997: 315)

- (50) ez a piros alma
this the red apple
'this red apple'

In Georgian adjectives commonly precede the noun, but may follow (in which case they are inflected like nouns; cf. section 1.5.4.1–2 on apposition):

Georgian (Fähnrich 1986: 53–4)

- (51) a. paṭara mt-eb-i b. mt-eb-i paṭar-eb-i
small mountain-PL-NOM mountain-PL-NOM small-PL-NOM
'the small mountains' 'the small mountains'

The adjective in Hittite appeared on either side of the noun and agreed in gender, number, and case (morphemes in capitals are Akkadian or Sumerian forms used to write Hittite; cf. e.g. Friedrich 1974: 22–3; Garrett 1990: 266):

Hittite (Luraghi 1990: 318)

- (52) GAL-in arunan
big-ACC.SG.C sea:ACC.SG.C
'the big sea'

Ika is an interesting case, in that most adjectives appear with the copula *kawa* 'seem'. If *kawa* indeed serves as a copula in this construction, we are dealing with a language in

which the majority of adjectives can only be used predicatively when modifying a noun.

Ika (Frank 1990: 32)

(53) kakarón aroma

shotgun_shell empty

'empty shotgun shell'

(54) paka aw Λn? kawa

cow big seem

'big cow'

Consider in this context also Nahali, in which attributive adjectives may be followed by *kā*, which is categorized as a copular element and probably a loan from Kurku (Kuiper 1962: 270; see also examples on pp. 292, 294):

The normal form of the copula is *kā*, just as in Kurku. Nahali also agrees with the latter language in using *kā* after adjectives, e.g. *awal kā kuprā* 'good cloth', like Kurku *awal kā lija*, id. 'Its force', says Drake, Grammar of the Kúrkú Language 12, 'is generally (p.126) emphatic and distinctive, but not seldom its retention or omission makes no appreciable difference in the sense'. However, Nahali differs from Kurku in that it apparently also uses *kā* after substantives to form adjectives, like Tamil *ana*.

Kayardild has a large set of adjectives, which normally precede the noun; like all adnominal modifiers they agree in case (Evans 1995: 235).

Kayardild (Evans 1995: 233)

(55) dathin-a jungarra dangka-a

that-NOM bie:NOM man-NOM

'that big man'

Ket adjectives invariably occur before the noun (Werner 1995: 198, 1997a: 323).

Ket (Werner, personal communication)

(56) kin 'e qo· aqta dε⁷ŋ

DEM ten good people

'these ten good people'

In Hmong Njua, the Hmong dialect that was used for this study, adjectives follow the noun:³⁴

Hmong Njua (Harriehausen 1990: 144)

(57) ob	phau	ntawv	loj	nua
	two	CLF	book	big DEM

Although adjectives in SE Asian languages are usually treated as a subclass of verbs (cf. Ratliff 1987: 82), I have categorized Hmong Njua as having a distinct class of adjectives for the simple reason that in this language verbal modifiers (as opposed to adjectives) must occur with a relative pronoun *kws* (Harriehausen 1990: 141, 224):

Hmong Njua (Harriehausen 1990: 141)

(58) Tug	miv	tub	kws	ua si	miv
	CLF	boy	REL	play	small

(59) Lub	ruuj	kws	kuv	yuav	kim
	CLF	table	REL	1SG	buy expensive

Adjectives in Nama Hottentot precede the noun in simple, integral NPs; in postnominal position they are in an appositional relationship with the rest of the (p.127) NP. Recall that the Nama NP ends in a person-number-gender suffix (section 3.3.2.2.3):³⁵

Nama (Hagman 1974: 65)

(60) káise	kái	'do-p
	very	big man-3SG.M

Nasioi adjectives can occur on either side of the noun, but they usually occur in postnominal position. In an indefinite NP the adjective takes a plural affix when modifying a human noun; compare:

Nasioi (Rausch 1912: 121, 123)

(61) kíri	námeu
breadfruit	ripe

(62) a.	manikuma	natáko	b.	maniku	natako-nukan
	woman	weak		woman	weak-PL

Since *-nukan* only appears with human referents, the adjective alone may suffice:

Nasioi (Rausch 1912: 121)

- (63) a. nanin norúkāmpe b. norúkāmpe-nukan
man intelligent intelligent-PL
'an intelligent man' 'intelligent people'

In the case of a definite singular NP the adjective agrees with the head noun in class and number:

Nasioi (Rausch 1912: 122)

- (64) bián-o petá-no
banana-CL6.SG ripe-CL6.SG
'the ripe banana'

But often the class suffix is only expressed on the (postnominal) adjective: Nasioi (Rausch 1912: 122)

- (65) a. koi-nave peté-nave b. koi peté-nave
tree-CL15.SG weak-CL15.SG tree weak-CL15.SG
'the soft tree' 'the soft tree'

(p.128) Since *-nave* is the class suffix for wood and wooden objects (Rausch 1912: 111), it is often enough to express the adjective plus class suffix:

Nasioi (Rausch 1912: 122–3)

- (66) peté-nave
soft-CL15.SG
'the soft tree'
- (67) natako-nani
weak-CL2.SG [Class 2 is used for women]
'the weak woman'

Definite plural NPs are marked by the suffix *-ni* or *-nanka*:

Nasioi (Rausch 1912: 123)

(68) a. maniku-ni natako-ni	b. maniku-nanka natako-nanka
woman-PL weak-PL	woman-PL weak-PL
'the weak women'	'the weak women'

An attributive adjective in Ngalakan normally follows the noun and is sometimes inflected to agree in noun class, case, and (rarely) number (usually only the head of the NP is case marked; M = masculine noun class prefix; *ṇu* is the short form and *nugu* the long form):

Ngalakan (Merlan 1983: 80–1)

(69) ḡu-go ⁷ je ḡugu-miṛpara gaña ⁷
M-that M-child small
'the/that small child'
(70) ḡu-go ⁷ je mungu-balku mu-gengeŋ

M-that MU-rope MU-long
'that long rope' [mu(ngu)- = class prefix]

Ngiti adjectives do not normally agree with the head noun, but three adjectives have suppletive forms for singular and plural number (Kutsch Lojenga 1994: 234): the equivalents of 'big', 'small', and 'other' (the latter I do not regard as a proper adjective; see section 7.6.1). Recall that only human nouns have plural forms:

Ngiti (Kutsch Lojenga 1994: 340, 345)

(71) ádzí imbi
long rope
'a long rope'

(72) a. ádrùngbă dza	b. ádròdró dza
big:SG house	big:PL house
'a big house'	'big houses'

(p.129)

(73) a. ádrùngbă ngba	b. ádròdró nzo
big:SG child	big:PL children
'a big child'	'big children'

Attributive adjectives follow the noun in Sumerian and Wambon. Occasionally the first syllable of the Wambon adjective is reduplicated to express some kind of plurality ('various', 'several'); cf. *ambunop* 'little' vs. *ambambunop* 'little (PL)'.

Sumerian (Thomsen 1984: 63)

(74) é	gal
house	big
'the big house'	

Wambon (de Vries 1989: 39)

(75) evo	kap	ambunopalín
that	man	very_little
'that very little man'		

Table 4.3 summarizes the syntactic properties of the adjectives in the languages mentioned in the above sections; cases of doubling (i.e. ANA; cf. note 48) are discussed in section 4.3.4.

Table 4.3. Adjective-noun order in languages of types 1–3: V/N/A, V-N/A, V-N-A

Abkhaz	ANA	Hmong	Njua	NA	Nama	AN
Alamblak	ANA	Hittite		ANA	Nasioi	ANA
Basque	ANA	Hungarian		AN	Ngalakan	NA
BD Creole	AN	Human		AN	Ngiti	AN
Bukiyip	ANA	Ika		NA	Quechua, Imb.	AN
Burushaski	AN	Kayardild		AN(A?)	Samoan	NA
Dutch	AN	Ket		AN	Sumerian	ANA
Georgian	ANA	Nahali		NA	Turkish	AN
Guaraní	NA				Wambon	NA

4.3.2. Languages of intermediate type 3/4: V-N(-A)

For a number of reasons the following languages have all been classified as intermediate type 3/4: Babungo, Bambara, Chukchi, Gude, Kisi, Oromo, Pipil, Sarcee, and Tamil.

First there are languages that have so few adjectives that it would be wrong to suggest that the languages in question have a major class of adjectives and should therefore be classified as belonging to type 3. On the other hand, since these languages do have a small number of adjectives it would also be wrong to classify them as languages without adjectives (type 4). As there is no principled way to determine in absolute numbers what constitutes a minor or major word class, I have decided more or less arbitrarily

that languages with up to two dozen (p.130) adjectives have a minor class of adjectives and will be classified as intermediate type 3/4. When the precise number of adjectives in a language was left unspecified I relied on the author's judgement. According to this criterion the following languages have a minor class of adjectives: Babungo, Bambara, Gude, Kisi, Oromo, and Tamil.

In Babungo nearly all adjectives are said to be derived from the progressive aspect form of the verb; only eleven forms cannot be related to verbal forms, at least not synchronically (Schaub 1985: 233–4, 245, 371–3).

Babungo (Schaub 1985: 77)

(76) *ŋkáw jéə kāŋ*

chair good my

'my good chair'

Adjectives are marked (by affix or tone) according to the noun class membership of the head noun they modify (Schaub 1985: 186, 234 f.).

Bambara only has a small number of underived adjectives (Brauner 1974: 38), which follow the noun and are marked for plural number with the suffix *-u* (recall that the plural suffix is rather similar to the third person plural pronoun and is probably better analyzed as a phrase-final clitic; cf. section 2.2):

Bambara (Brauner 1974: 39–40)

(77) *misira ba*

mosque big

'the big mosque'

(78) *se jan-u*

leg long-PL

'long legs'

A similar situation obtains in Tamil, where 'all but a very small handful of adjectival modifiers of nouns are derived forms. The set of those that cannot by simple rules be derived from noun or verb roots comprises such high-frequency items as *nalla* "good", *periya* "big", *cinna* "small", *putu* "new", *pazaya* "old" and a few basic color terms' (Asher 1982: 187). Tamil adjectives are invariable and only occur before the noun.

Tamil (Asher 1982: 62)

(79) *nalla manusan*

good man

'a good man'

Gude has the same number of adjectives as Babungo, namely eleven. They normally precede the noun, but when they follow they occur with the suffix *-kii*, which 'has an obvious resemblance to the definiteness markers' (Hoskison 1983: 53 f.); see also section 3.3.2.4.1).

(p.131) Gude (Hoskison 1983: 55)

(80) a. iirə la	b. la	irə-kii
old cow	cow	old-KII
'old cow'	'old cow'	

Other so-called adjectives either require the prefix *ma-* in prenominal position and both *ma-* and *-kii* in postnominal position (see also section 5.2.2.2 on the category of so-called 'adjectival nouns' in Gude). The meaning or function of the prefix, which also attaches to modifying verbs, remains unclear:

Gude (Hoskison 1983: 55)

(81) a. ma-ntə	la	b. la	ma-ntə-kii
MA-dead	cow	cow	MA-dead-KII
'dead cow'		'dead cow'	

Kisi has a closed class of twenty-three underived adjectives (Childs 1995: 127) which always follow the noun and which agree in class with the noun head (see also example (14) in Chapter 2).

Kisi (Childs 1995: 126)

(82) kpèlè-lé	nójéí-láŋ
bed-PRO	dirty-SUF
'a dirty bed'	

Oromo is classified as a language of intermediate type 3/4 because, according to Bender et al. (1976b: 145), '[as in other Ethiopian area languages], there is a shortage of adjectives ..., but this lack is made up for by the many verbs of state and becoming, by various phrases, and by the use of nouns in adjectival meanings'. Adjectives normally follow the noun, but in at least one dialect (Oromo of Wellega) they may appear in prenominal position 'for special foregrounding' (Gragg 1976: 191).³⁶

Pipil is classified as belonging to type 3/4 for two reasons. First, nearly all adjectives are borrowings; secondly, other adjectives (if they are that) are morphologically marked, but (as in the case of the *ma-* prefix in Gude above) the function or meaning of the marker remains unclear. The few morphologically unmarked adjectives that are not

Spanish loans are ‘in origin basically [members of] some other lexical category which have come to be employed as adjectives’ (Campbell 1985: 63–1):

Pipil (Campbell 1985: 120)

(83) se: siwa:pil hermosa

a girl pretty

‘a pretty girl’

(p.132) It seems that only Spanish loans may occur after the head noun (Campbell 1985: 108). In addition there are several other types of ‘adjectives’, such as the group of ‘basic adjectives’, which take the suffix *-k* after a root vowel and *-ti-k* after a consonant but which are ‘semantically unmarked or neutral’ (Campbell 1985: 61–2).

Pipil (Campbell 1985: 120)

(84) ne chih-chi:l-tik tsapu-t

the PL-red-TIK zapote-ABS

‘the red zapotes’

Other types are derived forms or the morphological marker is clearly related to verbal affixes that are used to express such notions as *inchoative* and *perfect*. Note finally that the adjective can be reduplicated to express plurality (Campbell 1985: 64).

In Chukchi an adjective can occur as a free or an incorporated form.³⁷ When it appears as a free modifier, it may appear before or after the noun and is preceded by the prefix *n(ə)-* and the suffix *-qin(e)* or *-qen(a)*; the incorporated form always occurs prenominally (Comrie 1981a: 251).

Chukchi (Koptevskaja-Tamm n.d.; examples originally from Skorik (1960–77: i. 421–9)

(85) a. n-ilgə-qin-Ø qoranə b. elgə-qoranə

STAT-white-3-SG deer:ABS.SG white-deer:ABS.SG

‘a white deer’ ‘a white deer’

Thus it seems that we can only speak of adjectives in Chukchi (as defined in Hengeveld’s terms) when they are incorporated (see also section 10.2.4.1).

Although there is little doubt that most ‘qualifiers’ or ‘adjectival modifiers’ in Sarcee have a verbal origin, their exact morpho-syntactic status is not quite clear (Cook 1984: 67):

The qualifiers are not free forms having a fixed position immediately after the noun, nor are they suffixes in the usual sense. For the purpose of discussion, however, the qualifiers are not considered an independent constituent within NP ... Two most frequently used qualifiers are *tsìtl'á* and *tcíw*, which are obviously related to verb stems *-tsòtl'á* ‘to be small’ and *-tców* ‘to be big’. Not every qualifier

is traceable to a verb stem. Nouns with a qualifier are different morphologically and semantically from nouns with a relative clause.

(p.133)

Table 4.4. Adjective-noun order in languages of intermediate type 3/4: V-N(-A)

Babungo	NA
Bambara	NA
Chukchi	ANA (AN order with incorporated A)
Gude	ANA
Kisi	NA
Oromo	ANA
Pipil	ANA
Sarcee	NA
Tamil	AN

The former is like a compound and the latter a phrase, comparable to the English nominal compound ‘bláckbírd’ and phrase ‘bláck bírd’.

Sarcee (Cook 1984: 68)

(86) t̪i-yáná
dog-old
‘old dog’

Table 4.4 gives an overview of the word order possibilities in languages of type 3/4.

4.3.3. Languages without a flexible or distinct class of adjectives

The following languages do not have a flexible or distinct class of adjectives: Burmese, Galela, Hixkaryana, Koasati, Korean, Krongo, Lango, Mandarin Chinese, Nivkh, Nung, Nunggubuyu, Vietnamese, Tsou, and West Greenlandic. These are all type 4 languages, which employ verbs and/or nouns to express adjectival notions, but it is also possible that a language employs appositional constituents because it has no integral noun phrase (see Nunggubuyu below).³⁸

(p.134) In his grammar of Lango, Noonan writes that this language lacks a distinct category of (attributive) adjectives; in fact, almost every modifier takes the form of an embedded structure (Noonan 1992: 103):³⁹

Associative constructions aside, there is no grammatical construction in Lango which is attributive and does not involve embedding. To translate an attributive adjective from English to Lango, it is necessary to resort to a relative clause construction, i.e. adjectives are always predicates in Lango. Compare [my numbering—JR]:

(87)	küll	à	ŋwé
	wart_hog	AP	3SG:smelly:HAB
‘a smelly wart hog’			

and

(88)	küll	à	òwòpé
	wart_hog	AP	3SG:follow:PERF:SG.O
‘that wart hog that followed me’			

The interpretation of

(89)	küll	ŋwé
	wart_hog	3SG:smelly:HAB
‘the wart hog is smelly’		

is only predicative. Thus, adjectives, are exactly like verbs in this respect since, verbs, too, require relative clause construction in order to modify nouns.

In Mandarin Chinese (as in many other Asian languages) adjectival notions are expressed by verbs, which, incidentally, also explains the absence of a copula in the Mandarin counterpart of sentences such as English ‘X is (very) pretty’. The modifying verb is followed by the ‘relativizer’ (Gao 1994: 494) or ‘nominalizer’ (Li and Thompson 1989: 118) *de*; compare:⁴⁰

Mandarin Chinese (Gao 1994: 494)

(90)	Mǎlì	hěn	piàoliàng
Mary very pretty			
‘Mary is very pretty’			

(91)	yī-wèi	hěn	piàoliàng	de	gūniang
one-CLF very pretty DE girl					
‘a very pretty girl’					

(p.135) Consider also the following remarks by Li and Thompson (1989: 826–7), but recall that in terms of the definitions provided in Chapter 1 only the third criterion is most relevant:

Strictly speaking, there is no class of words in Chinese that we can call ‘adjective’. That is, while there are certainly words which denote qualities or properties of entities, from a grammatical point of view it is difficult to distinguish ‘adjectives’ from ‘verbs’. First, in Chinese, words denoting qualities and properties do not occur with a copula as they do in Indo-European languages. Secondly, quality and

property words in Chinese are negated by the same particle *bù* as are verbs ... Thirdly, when an ‘adjective’ modifies a noun, it occurs with the same nominalising particle *de* as verb phrases do ... For these reasons, it is sensible to consider quality and property words in Chinese simply as a subclass of verbs, one which we might call ‘adjectival verbs’.

There is also a construction without *de*, but this seems to form a more or less lexicalized compound (Li and Thompson 1989: 119; cf. also e.g. Huang 1989; Chappell and Thompson 1992; McCawley 1992; Sackmann 1996a):

In general, adjectives that modify the noun without the particle *de* tend to be more closely knit with the noun. The consequence is that the adjective-plus-noun phrase tends to acquire the feature of being a *name* for a category of entities. The relative clause usage of adjectives, on the other hand, always has the function of further clarifying or delineating the reference of the head noun.

Compare:

Mandarin Chinese (Li and Thompson 1984: 118)

(92) a. hóng de huā	b. hóng huā
red DE flower	red flower
‘a flower that is red’	‘a red flower’

Perhaps the difference is best captured by comparing it again (see above on Sarcee) with English ‘black bird’, i.e. ‘a bird that is black’, and ‘blackbird’, i.e. ‘the kind of bird that belongs to the species “blackbird”’. Similar differences in meaning between the free modifier-noun construction and the compound form are attested in other classifier languages in the sample. In Vietnamese, for example, the difference is mainly indicated by different stress patterns (Löbel 1999: 280, 309). Thus, depending on the pronunciation, *áo dai* [dress be_long] is translated as ‘long dress’ (‘dress which is long’) or as ‘long-dress’ (‘kind of dress that is typical of Asian women’).⁴¹

Apparently ‘adjectival’ attributes in Burmese and Korean are all derived verbal forms. In Burmese they usually follow the noun, but they also occur before it (Okell 1969: 43, 81; Wheatley 1987: 851). In the following example *d-pya* is the derived form of the verb *pya* ‘be blue’ (Okell 1969: 47, 50, 79).

(p.136) Burmese (Okell 1969: 50)

(93) eìñci ā-pya (or with the order reversed: d-pya eìñci)
shirt [ā-]be_blue
‘blue shirt’

In Korean adjectival notions are expressed by verbs in their ‘modifier form’, which means that they appear with the attributive (realis) suffix—(*u*)*n*. Compare:⁴²

Korean (Martin and Lee 1969: 202–3)

(94) khun [=khe + un] cip

be_large [=be_large + UN] house

'large house, house which is large'

(95) mek-un sālam

eat-UN person

'person who ate, person who has eaten'

Korean (Lee 1993: 84)

(96) ôce chac-a o-n khi-ka

yesterday visit-CN come-(U)N height-NOM

cak-ûn salam-i

(be_)small-UN person-NOM

'the short man who visited us yesterday'

In Vietnamese, too, verbal elements are used to express adjectival notions in that the modifiers in question behave the same as descriptive or stative verbs (cf. Emeneau 1951: 85; Thompson 1965: 217; Honey 1972: 278; Nguyễn Đình Hoà 1987: 785, 790; 1997: 119; Löbel 1999: 285); sometimes the modifying phrase is introduced by the conjunction *mà*.

Vietnamese (Thompson 1965: 124)

(97) con chó nhỏ này

CLF dog be_small this

'this small dog'

Vietnamese (Nguyễn Đình Hoà 1997: 174)

(98) con dao (mà) anh cho tôi muốn

CLF knife (MA) you give me borrow

'the knife you lent me'

Nung adjectives do not seem to behave differently from attributive or predicative verbs either; hence Saul and Freiberger Wilson (1980: 32) treat them as 'simple, (p.137) reduplicated or complex descriptive verbs':

Nung (Saul and Freiberger Wilson 1980: 33)

(99) vahng lōng

boy be_big

'the big boy'

Nung (Saul and Freiberger Wilson 1980: 69)

(100) slóng óhng déhc kíhn hô tê
two CLF child eat beg those
'those two begging children'

Nung (Saul and Freiberger Wilson 1980: 16)

(101) vahng khị tú bê tê
boy ride CLF goat that
'that boy (who was) riding the goat'

Nung (Saul and Freiberger Wilson 1980: 79)

(102) ahn hòn cáu lẹn khàu pay tê
CLF house I run into go that
'the house I was running into'

Sometimes the relative marker *ti* or *ti-vä* 'who, which' is used:

Nung (Saul and Freiberger Wilson 1980: 78)

(103) kê (ti-vä) tóhc cộn thíhn tê ni
man (who) lost CLF stone that FOC
'the man who had lost the stone'

We have already seen in section 1.5.1 that in Galela the first syllable of the verb is reduplicated (=RED) to produce a participle, which can then be used as a noun modifier, compare:⁴³

Galela (van Baarda 1908: 36)

(104) o tuhu °da lòha
ART house 3 be_beautiful
'the house is beautiful'

(105) o tuhu °da lo-lòha

(p.138) Hixkaryana is simply stated not to have adjectives ('there are no adjectives or adjective-noun phrases'—Derbyshire 1979: 131); in fact, there does not seem to be any other kind of modifier construction in the NP to specify qualitative properties. Apparently such properties are only expressed in a non-verbal predicative construction ('[the equative construction] handles what in other languages are adjective-noun and demonstrative-noun relations'—Derbyshire 1979: 37).

Kimball (1991: 479) states that in Koasati 'the adjectival phrase, if it may be so called, is a nominalized verb phrase'. Thus, the adjectival phrase in (106b) is derived from (106a):⁴⁵

Koasati (Kimball 1991: 480)

(106) a.	í:sak	hatkáhcq	b.	í:sa	hátka
	í:sa-k	hátka-Vhco-ɍ		house	one_which_is_white
	house-S	be_white-HAB-PTM		'a white house'	
'the house is white'					

The only difference between predicative and attributive verbs in Krongo is that verbs that are used predicatively precede the noun whereas attributive verbs follow and must be prefixed with a 'connector' (CN), which only appears in the so-called narrative construction on predicative verbs (for more details, see Reh 1985: 193 f., 250 f.; notice that, due to fusion of the linker morpheme *ŋ-* with the neuter agreement prefix *n-*, the resulting form is identical with the neuter prefix: *n-*):

Krongo (Reh 1985: 251)

(107) a.	n-òoróobó	nì	b.	nì	n-òoróobó
	N-IMPF.be_evil	snake		snake	CN-IMPF.be_evil
	'the snake is evil'			'an evil snake'	

What information there is about Nivkh suggests that adjectival notions are expressed by verbal roots that however do not occur as free adnominal modifiers but are part of a poly synthetic construction (Gruzdeva 1998: 16, for references and discussion, see Mattissen and Drossard 1998: 38–44, 50).⁴⁶

Nivkh (Mattissen and Drossard 1998: 51; taken from Jakobson 1971c: 80)

(108)	hun-tleulan-̪i̪ñ
	that-be_white-hill
	'that white hill'

Nunggubuyu is classified as a type 4 language, mainly on the basis of statements provided by the author. Although Heath (1984: 498) first adopts the fuzzy-NP analysis in his discussion of Nunggubuyu NP structure (meaning that he does not (p.139) want to say that NPs are either appositional words or ‘tightly-knit’ syntactic units), he later writes:

In general, inspection of a large number of examples suggests that combinations of a (semantically) descriptive NAdj [+adjectival noun—JR] and a nuclear (referential) noun ... do not differ in surface structure from appositional sequences ... NAdj ... are formally autonomous nouns which may be closely juxtaposed to a nuclear noun but are best viewed as separate, appositional units.

Like some other isolating languages, Tsou represents a difficult case when it comes to characterizing its parts-of-speech system in terms of Hengeveld’s definitions. Due to the absence of unambiguous morpho-syntactic criteria in Tsou, it is practically impossible to decide whether noun modifiers are simply unmarked verbs or constitute a distinct class of adjectives. From the available data it seems that Tsou lexemes which express adjectival notions do not require special measures when they are used as noun modifiers (which suggests they are adjectives in Hengeveld’s sense), but neither do verbs (cf. Szakos 1994: 78):⁴⁷

Tsou (Szakos 1994: 79, 136)

(109) tpitpiei ci eatatiskova
deceive CI person/people
‘deceitful person’

(110) kua 'onja ci zomi
black CI bird
‘a black bird’

Note that any prenominal modifier is separated from the head noun by the attributive particle *ci* (Szakos 1994: 162):

Tsou (Stanley Starosta and Elizabeth Zeitoun, personal communication)

(111) yuso ci meoi ci av?u tan?e
two CI big CI dog this
‘these two big dogs’

Finally, in West Greenlandic, according to Fortescue (1984: 204), ‘adjectival function is served by a subset of (stative) verbal and nominal bases (notably including intransitive participial forms of the former in attributive position)’.

West Greenlandic (Fortescue 1984: 108)

(112) illuqarvik miki-suu-suq

town small-be-INTR.PART

'a small town'

(p.140) 4.3.4. Doubling of adjectives⁴⁸

In the following languages some or all adjectives may appear on either side of the head noun: Abkhaz, Alamblak, Basque, Bukiyp, Georgian, Hittite, Kayardild, Nama Hottentot, Nasioi, Ngalakan, Sumerian (see section 4.3.2 on doubling in the languages of type 3/4: Chukchi, Gude, Oromo, and Pipil).⁴⁹

In Abkhaz only adjectives expressing nationality precede the head noun (Hewitt 1979: 56), in Basque no more than a few adjectives, including e.g. 'american', are said to appear before the head noun (Saltarelli 1988: 75), and in Modern Georgian placing the adjective after the noun is 'one of the features of archaic style' (Testelec 1997a: 249; cf. section 1.5.4.1).

I have already mentioned that in Alamblak adjectives relating to the notion 'value' often occur in postnominal position, whereas those relating to the notions 'physical property' and 'age/human propensity' tend to precede the noun. Color adjectives always precede the noun and adjectives of dimension do not seem to have a preferred position (Bruce 1984: 78, 99).

In Nama Hottentot only prenominal adjectives are part of the integral NP; (Chapter 1); postnominal occurrences of the adjective must be regarded as appositional constructions. In the case of Gude we may also be dealing with appositional modifiers, since postnominal adjectives (if they are that) may only follow the noun, if they are provided with a suffix that resembles the definiteness marker (Hoskison 1983: 53).

Not much is said about the postnominal occurrence of adjectives in the Australian languages Kayardild and Ngalakan. Evans (1995: 235) writes that as a rule Kayardild adnominal modifiers precede the noun, 'except that one modifier may be postposed'; presumably this also includes members of the adjective class. In a similar vein Merlan (1983: 81) states that an adjective in Ngalakan 'almost always follows the noun ... but exceptions are found'. Recall, however, that she (*ibid.* 83) also argues that constituents of the Ngalakan NP could be analyzed as appositional elements (section 1.5.4.2).⁵⁰

It is not stated which factors are involved in the placement of adjectives in Hittite and Nasioi, except that in Nasioi their most natural position appears to be after the noun (Rausch 1912: 121; see above for more details). It is not entirely clear what determines the position of the adjective relative to the noun in Bukiyp (p.141) either, but one gets the impression that the prenominal position is more common since Conrad (1991: 57) writes that adjectives (as well as demonstratives and numerals) may follow the noun if they are 'the only optional constituents occurring in the [noun] phrase'. Finally, in Sumerian only one adjective may appear before the noun, *kug* 'holy, pure') but only in 'divine names' (Thomsen 1984: 64).

4.3.5. When can a language have adjectives?

Data in the sample languages indicate that the occurrence of adjectives in a language depends on a lexical property of the nouns in that language (as before, in this context the category of ‘nouns’ is restricted to nouns used in connection with a single, discrete spatial entity in the real world).

Let us return to Table 1.3 (=Table 4.5), which lists the languages according to the parts-of-speech systems as defined by Hengeveld (1992b); the italicized languages all employ sortal (or numeral) classifiers: Hmong Njua, Burmese, Korean, Mandarin Chinese, Nivkh, Nung, Vietnamese.⁵¹

What is noteworthy here is that, apart from Hmong Njua, all classifier languages belong to type 4, which has only two distinct word classes: verbs and nouns, but no adjectives. In other words:

- (113) If a language has classifiers then it usually has no adjectives (or: as a rule a language only has adjectives if nouns are in a direct construction with a numeral).

Interestingly this statistical universal can be turned into an absolute implication, if we take into account the lexical semantics of the noun. I argued in Chapter 3 that the employment of any kind of classifier (whether sortal, mensural, or general) is normally a good indication that the modified noun belongs to one of the –Shape nominal subcategories (sort noun, mass noun, general noun), i.e. nouns which do not include as part of their lexical meaning the notion of spatial boundedness or discreteness. Since only discrete entities can be counted or measured, a special construction is needed, namely the classifier phrase, to count or measure referents that have been construed on the basis of a –Shape noun. Here we are mostly interested in sortal classifiers (the italicized languages in Table 4.5 are all languages that employ sortal classifiers rather than general classifiers), and we also saw that—due to grammaticalization processes—sortal classifiers may assume other functions. This has also happened in Hmong (cf. Bisang 1996, 1999), the only language in the sample that has both classifiers and a distinct class of adjectives, and apparently classifiers are now (also) used to specify that the referent is a singular object or a collective, i.e. a singleton set or a collective set. In other words, since Hmong Njua classifiers do the same thing as (p.142)

Table 4.5. Parts-of-speech systems of languages in the sample

Type 1	V/N/A	Samoan
Type 2	V- N/A	Human, Imbabura Quechua, Turkish
Type 3	V-N- A	Abkhaz, Alambalak, Basque, Berbice Dutch Creole, Bukiyp, Burushaski, Dutch, Georgian, Guaraní, Hittite, Hmong Njua, Hungarian, Ika, Kayardild, Ket, Nama Hottentot, Nasioi, Ngalakan, Ngiti, Sumerian, Wambon
Type 3/4	V-N(- A)	Babungo, Bambara, Chukchi, Gude, Kisi, Oromo, Pipil, Sarcee, Tamil
Type 4	V-N	Burmese, Galela, Hixkaryana, Koasati, Korean, Krongo, Lango, Mandarin Chinese, Nivkh, Nung, Nunggubuyu, Tsou, Vietnamese, West Greenlandic
Type 4/5	V(-N)	Cayuga
Type unknown		Etruscan, Meroitic, Nahali

the nominal aspect markers in a typical set noun language (such as Oromo), nouns in Hmong Njua are better categorized as set nouns. Compare the following examples again (cf. section 4.2.1.2):

Hmong Njua (Harriehausen 1990: 115, 117)

(114) a.	kuv	yuav	tsev			
	1SG	buy	house			
	'I buy a house/(some) houses'					
b.	kuv	yuav	lub	tsev		
	1SG	buy	CLF	house		
	'I buy the house'					
c.	kuv	yuav	cov	tsev		
	1SG	buy	CLF	house [c<9v == group classifier]		
	'I buy (the) houses'					
(115)	kuv	pum	peb	tug/*cov	tlev	luv
	1SG	see	four	CLF/*COV	dog	big
	'I see four big dogs'					

These examples demonstrate that Hmong Njua nouns have the same properties as nouns in Oromo and other languages with set nouns. The first example (*a*) shows that Hmong Njua nouns are (still) transnumeral: the referent of an NP headed by an unmarked head noun may involve one or more individuals or objects. The second example (*b*) illustrates how the classifier (without the numeral) specifies that the (definite) NP has singular reference (singleton set). The third example (*c*) shows how the (erstwhile) collective or group classifier *cov* indicates that the NP refers to a collective entity (collective set), and the last example demonstrates that the collective aspect marker and the numeral are mutually exclusive (section 2.2.2).

(p.143) Together these examples suggest that Hmong Njua classifiers have grammaticalized and that Hmong Njua employs set nouns [+Shape] rather than sort nouns [-Shape]. This means, of course, that we now have an exceptionless universal, because all languages of type 3 use +Shape nouns (section 2.4):

(116) If a language has a distinct class of adjectives, then the nouns in that language are generally characterized by the feature [+Shape].

This gives us three possible combinations, each of which is attested in the sample:

1. [A and B]—languages with a distinct class of adjectives and first order nouns that are generally characterized by the feature [+Shape]; i.e. the language uses singular object nouns and/or set nouns, neither of which require a classifier when modified by a numeral. In the sample this involves all the languages of type 3; for example Hungarian (Moravcsik 1997: 315).

(117)	a.	ez	a	pirosA	alma	b.	ez	a	két	alma
		this	the	red	apple		this	the	two	apple
	'this red apple'					'these two apples'				

2. [not A and B]—languages without a distinct class of adjectives but with [+Shape] nouns (such as singular object nouns and/or set nouns); these are basically the languages that do not belong to type 3 in Table 4.5 above and which do not employ (sortal or general) classifiers: Samoan, Turkish, Hurrian, Imbabura Quechua, Galela, Hixkaryana, Koasati, Krongo, Nunggubuyu, Tsou. Krongo (Reh 1985: 252)

(118)	a.	nóo-còorì	nk-íifi
		PL-house	CN.PL-IMPF:be_big
	'big houses'		
	b.	nóo-còorì	nk-óotònò
		PL-house	CN.PL-IMPF:be_three
	'three houses'		

It does not make a difference whether languages of type 3/4 are regarded as type 3 or as type 4 languages. Since none of the languages in this intermediate category employs sortal classifiers they would not violate the implication proposed in (116). The following example is from Bambara (which only has a minor class of adjectives); *tími* 'be sweet' is one of the stative verbs that must take the suffix *-man* before it can be used as a noun attribute:
Bambara (Kastenholz 1989: 31. 79)

(119)	mángoro	tími/ stative-man
	mango	be_sweet-MAN
	'a sweet mango'	

(p.144)

(120)	sò	fia
	horse	two
	'two horses'	

3. [not A and not B]—languages without a distinct class of adjectives and without nouns that are characterized by the feature [+Shape]; these are basically the classifier languages (Hmong Njua being the exceptional case, of course): Burmese, Korean, Mandarin Chinese, Nung, Vietnamese.
Korean (Martin and Lee 1969: 202)

(121)	khun	[=khe + un]	cip
	be_large	[=be_large + UN]	house
	'large house, house which is large'		

(122)	kāy	twū	mali
-------	-----	-----	------

dog	two	CLF
	‘two dogs’	

The following type is logically excluded: *[A and not B], i.e. languages with a major (distinct) class of adjectives but without nouns that are characterized by the feature [+Shape]. Hmong Njua was the only potential counter-example in the sample: it has adjectives and uses classifiers, which typically occur in languages with sort nouns [–Shape]. But as I have argued above, Hmong Njua is the only classifier language in the sample that on a more or less regular basis indicates whether the referent is a collective or a singleton *set* entity (rather than a sort entity). Although it is probably true that Mandarin can also express the grammatical notion of collectivity with nominal NPs, the element in question (i.e. the suffix *-men*) is rarely used and seems to be restricted to non-monosyllabic human nouns (Li and Thompson 1989: 40). Note, incidentally, that the existence of languages with set nouns and without a major class of adjectives does not violate the implication formulated above.

Why is it that a language can only have a major class of adjectives if first order nouns such as ‘book’ and ‘girl’ have the lexical feature [+Shape]? I can only attempt to provide a tentative answer here.

We saw in Chapter 3 that languages may differ with respect to the encoding of ontological properties in the lexical information of a noun (notably regarding the features Shape and Homogeneity) and this is, of course, precisely the reason why different kinds of nouns (general noun, sort noun, set noun, singular object noun) can be used for the same object in the non-linguistic world. So in the case of nouns languages have a choice: either they have nouns whose lexical features ‘agree’ with certain ontological facts (notably [+Shape]) about the real-world correlate of the referent, or they have nouns whose lexical semantics do not mirror these ontological details.

In the case of adjectives, however, the choice between [+Shape] and [–Shape] does not seem to be available, simply because there is nothing in the (p.145) physical world to suggest that adjectival notions such as ‘poor’, ‘ripe’, or ‘green’ have (by themselves) a definite spatial outline: one can draw a picture of a house and even a waterfall, but one cannot draw a picture of ‘poor’ or ‘ripe’. In other words, properties designated by adjectives are all necessarily characterized by the feature [–Shape].

If it is true that adjectives are characterized by the feature [–Shape], then the reason why they only occur in languages with [+Shape] nouns may have to do with the fact that there is no good way to distinguish adjectives from nouns in languages using [–Shape] nouns (general nouns, sort nouns) in relation to (real-world) discrete objects. Conversely, one could hypothesize that only if a language uses [+Shape] nouns (like singular object nouns and set nouns) can it accommodate another major word class whose members are all exclusively characterized by the feature [–Shape]: adjectives (cf. Rijkhoff 2000, 2001).

4.4. Conclusion

In this chapter we have seen that qualitative properties of a referent may be indicated by grammatical elements (quality operators), namely individual and collective aspect

markers, which specify how the property that is designated by the head noun is expressed in the spatial dimension, and by lexical elements (quality satellites). Here I restricted myself to one particular kind of quality satellite, namely adjectives, the reason being that if a language has adjectives, it is these elements that are typically employed as such. In languages that have no distinct class of adjectives other lexical means are normally used to express more or less inherent properties of the referent (thus the English adjective ‘kind’ would be expressed as ‘... (with) kindness’ or ‘... who is kind’, i.e. as a qualifying adnominal NP or relative clause). Finally I put forward an implicational universal, stating that a language can only have adjectives if it employs singular object and set nouns (i.e. nouns that are characterized by the lexical feature +Shape). In the next chapter I will be concerned with quantitative properties of referents.

Notes:

(1) I am aware of at least two authors who used the phrase ‘nominal aspect’ before me (cf. Rijkhoff 1991), albeit in a different sense, namely Benzing (1955: 57) and Heath (1984: 172); see also Meisterfeld (1998). Compare in this context also Leech (1969); Greenberg (1972: 30); Seuren (1974b: 4); Borg (1980), Cusic (1981: 1–2); Mourelatos (1981: 202 f.); Dahl (1985); Langacker (1987); Krifka (1989); Jackendoff (1990: 29); Brinton (1991), Lucy (1992: 74); see also Dik (1997: 138, 163–6) and Van Valin and La Polla (1997: 56–7).

(2) See Kuhn (1982) about lexical as well as grammatical expressions of the notions ‘collective’ and—to a lesser extent—‘singulative’ (cf. Greenberg 1972; Allan 1976; Biermann 1982; Anderson 1985: 174). I will also ignore lexically derived collective nouns, but note that there is a difference between derived collectives on the one hand and set nouns with collective aspect marking on the other in that the former (just like basic collective nouns) but not the latter can occur in the plural; see e.g. Saltarelli (1988: 198–9) on Basque; Donaldson (1980: 102) on Ngiyambaa.

(3) As will be demonstrated below, there is some variation with respect to the occurrence of the so-called number marker on non-numerated set nouns. In some languages it is compulsory (e.g. Hungarian), in others it is optional (Oromo), and there are also languages in which it is totally absent (e.g. most nouns in Ngiti).

(4) Singulative markers are regarded as one of the features of the Ethiopian Language Area (Ferguson 1976: 74; Unseth 1988: 88). The same marker has also been called ‘particularizing determiner’, ‘individualis’, ‘suffisso d’unità’, ‘forma individuante’ (Tucker and Bryan 1966: 525). See also Newman (1990: 133–4) on singulative marking in the Chadic languages: ‘In some present-day languages (e.g. Masa) the grammatically masculine singular form may be semantically collective or generic Such generic words are then made singulative or individuative by changing the gender to feminine.’ Note that singularity is also explicitly expressed in the Bantu-type noun class affix and by a numeral classifier in certain languages of SE Asia (see Chapter 5).

(5) See e.g. Holmer (1947: 61, 86 f.) on Cuna; Churchward (1941: 15–16) and Dixon (1988: 175) on Fijian; Beekes (1990: 213) on Proto Indo-European; Derbyshire (1979: 126) on Hixkaryana; Martin and Lee (1969: 32) on Korean; Chao (1961: 40) on Mandarin Chinese; Merlan (1982: 86) on Mangarayi; Du Feu (1996: 135) on Rapanui; Cook (1984: 65) on Sarcee; Thomsen (1984: 59) on Sumerian; Swanton (1911) on Tlingit; Benzing (1959: 722) on Chuvash; Szakos (1994: 68) on Tsou.

(6) For a more extensive discussion of verbal aspect I refer the reader to Verkuyl (1972, 1993); Anderson (1973); Comrie (1976); Lyons (1977: chapter 15.6); Hopper (1982); Tedeschi and Zaenen (1981); de Groot and Tommola (1984); Chung and Timberlake (1985); Thelin (1990); Bybee (1985); Verkuyl (1993); Bybee et al. (1994); Sasse (2001). See Dik (1997: 221 f.) for a discussion of verbal aspect within the theoretical framework of Functional Grammar.

(7) My data suggest that collective aspect is overtly indicated more often than singulative aspect. Interestingly, collective aspect is also the aspect that is marked for both Shape and Homogeneity (+Shape and +Homogeneity). It would be interesting to find out if the same cross-linguistic correlation obtains for perfective aspect, the *verbal* aspect that is marked for either aspectual feature (+Beginning and +Ending); for some positive evidence see e.g. Bybee and Dahl (1989: 95). Note that there is also a diachronic relation between collective and perfective marking. Von Garnier (1909) has already argued that the collective marker has developed into a perfective marker in several Indo-European languages.

(8) In many languages number agreement is not a relevant parameter, either because there is no number agreement at all between one or more arguments and the verb (e.g. Berbice Dutch, Ngiyi), or because number distinctions are not (or not always) expressed in the third person (Galela, Imbabura Quechua).

(9) I am grateful to Ulrich Lüders (referring to Rebuschi 1990) who has pointed out to me that Basque *d-i-r(a)* may be glossed differently. As to the prefix *d-*, he says that its status is not clear: it has been analyzed as a person marker, a present tense marker, and as both; that *-i-* is possibly the intransitive root of *izan* ‘to be’, and that *-r(a)* is an irregular plural number marker which can only be found in intransitive forms referring to subject NP. Furthermore, he claims that ‘the corresponding singular form of *dira* is *da*, showing that *d-* is not subcategorized for number. It is neither 3SG nor 3PL.’

(10) Harris (1981: 22) adds that ‘although the dichotomy is required by prescriptive norms, the rule is not always followed by speakers’. Georgian has another plural suffix (*-n*), which, however, seems to be getting out of use. Interestingly, nouns with the +plural do not trigger ‘number discord’ in the verbal complex (cf. Fähnrich 1986: 158–9).

(11) One of the characteristics that is frequently encountered in Papuan languages is, according to Wurm (1982: 36), ‘the rarity of morphologically marked non-singular forms of nouns (this is however not true of members of the Torricelli Phylum, of some members of the Sepik-Ramu Phylum, and of some other Papuan languages)’.

(12) For discussion of number marking (if it is that) in North American languages, see e.g. Mithun (1988: 212; see also Mithun 1999: 79–94), who writes: ‘[In fact], in the majority of North American languages, nouns referring to nonhumans are not marked for number at all. Yet even when plurals do exist, they are not necessarily used mismatches [i.e. cases of ‘number’ discord—JR] are usually not the result of sloppiness on the part of the speakers, although this possibility has sometimes been suggested by frustrated grammarians.’

(13) Besides the particle *komo* ‘collective’, which is the usual collective marker with nouns, there is also the suffix *-yamo* ‘collective’, which is used mainly with a subclass of

derived nouns (Derbyshire 1979: 126).

(14) As for American Indian languages in general, Boas (1911a: 37–8) wrote (cf. also Sapir and Swadesh 1946 and note 12): ‘It would seem that, on the whole, American languages are rather indifferent in regard to the clear expression of plurality, but that they tend to express much more rigidly the ideas of collectivity or distribution. Thus the Kwakiutl, who are rather indifferent to the expression of plurality, are very particular in denoting whether the objects spoken of are distributed here or there. When this is the case, the distribution is carefully expressed. In the same way, when speaking of fish, they express by the same term a single fish and a quantity of fish. When, however, they desire to say that these fish belong to different species, a distributive form expressing this idea is made use of. A similar indifference to the idea of singular and plural may be observed in the pronouns of several languages.’

(15) See also Dixon (1980: 267) on ‘number’ in Australian languages: ‘In English each noun must be marked for number, but in Australian languages whether reference is to “one” or “more than one” need not be stated. Explicit number specifications can be made if required: plural (indicating three or more) is usually shown by reduplication ... dual can often be shown by an affix or by a separate number adjective “two”; singularity can be stressed by adding a number adjective “one” (agreeing with the head noun in case).’

(16) Almost one and a half centuries ago Castrén (1858: 16) wrote that the plural marker is often omitted in Ket.

(17) Notice that here plurality is expressed in the suppletion of the verb form (Kimball 1991: 323; cf. also Mithun 1999: 83).

(18) The form of the definiteness marker is determined by the class of the noun; *-nàva* is the suffix for trees and other wooden entities. Nasioi has two plural suffixes: *-ni* is used for small numbers and *-nanka* for large numbers (Rausch 1912: 114).

(19) Interestingly, non-human nouns often occur with a dual suffix (Heath 1984: 195), which seems to suggest that the dual marker is not a number marker but rather indicates that the referent consists of a two-member set rather than two individual objects. In this context it is perhaps interesting to note Mithun (1999: 93) writes that in the language of native North America collective markers often descend diachronically from duals (and reciprocals).

(20) The Tsou suffix *-si* serves both individuality and possession (Szakos 1994: 68).

(21) According to a description provided by the Miao Language Team (1972: 14), Hmong classifiers can be used for both definite and indefinite reference: ‘Each regular classifier has five forms, indicating the size, appearance, and definiteness of an object. Note, for example, the changes in lu 55 “clf. for round or hollow objects” [numbers correspond to tone values-JR]: lu 55 (definite or indefinite, large attractive), lai 55 (definite, ordinary), la 55 (definite, tiny), lai 35 (indefinite, ordinary), la 35 (indefinite, tiny).’

(22) See also e.g. Bisang (1996). Although I regard *cov* as a marker of collectivity rather than simple plurality, the glosses are those provided by Harriehausen (1990). One could hypothesize (as Ratliff indeed does; cf. Ratliff 1991: 696, 699) that ultimately *cov* may

become a real plural marker. Indeed, it has been established for a wide variety of languages that (besides e.g. gender markers and demonstratives; Frajzyngier 1997) collective markers are a common source for (real) plural markers. The diachronic relation between collective and plural markers has been observed in e.g. Kartvelian (Tuite 1992: 271), Ket (Werner 1994: 51), the Mesoamerican languages (Suárez 1983: 86), and the Semitic languages (Kuryłowicz 1976); cf. on the same phenomenon also Kuryłowicz (1964, 1972); Jespersen (1924: 195); Meillet (1967: 66); Comrie (1981a: 167); Menges (1968: 111–12).

This suggests one could propose the following chain of grammaticalization: collective noun/group classifier → collective marker → plural marker. Since classifiers occur with sort nouns, collective aspect is marked on set nouns, and plural marking is a typical feature of singular object nouns, we may also hypothesize the following development of noun types: (general noun+) sort noun → set noun → singular object noun.

(23) See also Anderson (1985: 174): ‘In Tlingit ... there is not really a category of plural, but rather of collective: the difference between a singular and a coherent group In fact, we can make a “plural” form *yuyai tlang* “the big whale”, which refers to a single whale but stresses that it was very large and had many parts to be cut out.’

(24) Cf. also Hetzron (1967: 194) on Agaw (or Awngi), another Cushitic language of the Afro-Asiatic phylum.

(25) Glosses suggest that the diminutive suffix -(o)si also makes the nouns singular (Kimball 1991: 449): *ná:nosi* ‘boy’, *tayyosi* ‘girl’. Although this is explicitly mentioned in Kimball’s unpublished dissertation (1985: 393, 394), it seems to have been omitted in the published version (Kimball 1991). See Jurafsky (1996) on the semantics of diminutives in general.

(26) In Dutch the diminutive is obligatory with certain derived nominals (cf. also Jurafsky 1996: 555 f. on the individuating sense of diminutives). For instance, there is a group of numerals which may be used to refer to banknotes which represent the amount of money corresponding to those numerals. Thus, a 10-guilder banknote is referred to as ‘een tien-tje’ (a ten-Dim), and a 100-guilder banknote ‘een honderd-je’ (a hundred-Dim). In other languages there seems to be a relation between the agentive affix and the marking or singularity or individuality. For instance in several languages in the sample (e.g. Hurrian, possibly Koasati) the singulative affix is more or less identical with the agentive marker (see also Greenberg 1981: 108, 1991a). For examples outside the sample, see e.g. Holmer (1946) on Cuna; Tucker and Bryan (1966: 526) on Saho; and Mous (1993: 63) on Iraqw (both Cushitic languages); cf. also English ‘a fiver’, ‘a tenner’.

(27) The same situation seems to hold true in e.g. Ewe (Westermann 1930: 47) and Shilluk (Westermann 1912:46); cf. also Lehmann (1982a: 60) and Frajzyngier (1997:211–14). Kinkade (1995: 348, 354) writes that one of the so-called plural suffixes in Upper Chehalis (Salishan, Amerind) has a collective meaning, is clearly related to third person plural forms, and derives from a lexical suffix meaning ‘person, people’ (note also that pluralization is optional and commonly absent—ibid. 359).

(28) See also DeBose (1974) on Papiamento and Langacker (1977: 80) on the hypothesized diachronic relation between diminutive and ‘number’ marking in Uto-Aztecán languages.

(29) See also Newman (1990: 18) on the employment of an erstwhile deictic element as a plural marker in the Chadic languages. Frajzyngier (1977: 55) suggests that in the Chadic languages masculine and feminine markers may have been transformed into plural markers. Other possible sources for the singulative aspect marker may be ‘one’ (cf. Lorimer 1935–8: i. 47 f. on the so-called singular suffix in Burushaski; see also Greenberg 1991a: 310 and Holmer 1946 on Cuna *-kwa*).

(30) Consider also this remark by Rowlands (1959: 38) on the Gambian Mandinka ‘plural’ suffix *-lu*: ‘This plural suffix has some characteristics which could justify writing it as a separate word. These are: (i) it may be regarded as having a separate accent of its own, since in falling intonation, unlike accentless words, it is a step lower in pitch than the preceding syllable of the O-suffix form ... , (ii) a few instances have been found in texts in which *-lu* is separated from its Nominal by the modifier *le* [probably a focus marker].’

(31) Cf. Bybee (1985: 5); Kuhn (1982); Anderson (1985: 162).

(32) Speiser (1941: 71–2) also wrote that at an earlier stage Hurrian was (in our terminology, of course) a type-1 language: ‘It follows that the Hurrian root was not at all times inherently noun, verb, or particle; but each of these categories was typified by its ability or inability to take on given grammatical markers, depending on the required function as part of speech which it was to fill in the sentence.’

(33) Many adjectives in Burushaski seem to be erstwhile verbal forms, since Berger (1998: 77) notes that quite a few adjectives end in the ‘ablative and participial suffix -um/-úm’ but that the original stem to which this suffix was attached no longer exists.

(34) In White Hmong, for example, there is a handful of adjectives that precede the noun (Ratliff 1991: 695). Notice also that in the description of the Wei Ning dialect of Hmong in Kweichow the adjective is sometimes preceded by an ‘adjective marker’ (here in boldface; numbers indicate tones): lu 55 li 33 p'y 55 ku 11 lo 55 [CLF bottle ADJ.MARKER big] ‘a big bottle’ (Wang 1972: 125).

(35) Nama Hottentot has a closed class of adjectives, but the size of the class is not specified: ‘Nama has a rather limited number of monomorphemic adjectives but these “simple adjectives” are frequently used’ (Hagman 1974: 59).

(36) See also Sasse (1976: 206) on Dasenech, another Eastern Cushitic language: ‘The number of radical adjectives in Dasenech is very limited. Many of those relations that we express by adjectives are rendered in Dasenech by relative constructions, genitives, or the like. All these expressions, as well as other “real” adjectives, are followed by the subordinator -ka.’

(37) The various sources give, however, different accounts as to the way adjectival notions can be expressed. According to Comrie (1981a: 251) incorporation used to be compulsory in the oblique cases but ‘in present-day Chukchi this rule is not observed rigidly, and it is possible to find oblique case noun phrases where the adjective agrees in case with its head rather than being incorporated’. Vladimir Nedjalkov (personal communication) states that the free variant is only possible in the absolute and Kämpfe and Volodin (1995: 37, see also p. 101) write that adjectival roots are never expressed as free modifiers: ‘Die Adjektivwurzeln des Tschuktschischen ... kommen

ausschließlich als gebundene Morpheme vor, nämlich als Bestandteile von Inkorporativkomplexen und in Prädikativ- und Komparationsformen.'

(38) Cayuga is the only type 5 or rather type 4/5 language in the sample (section 1.5.1). Mithun (1992: 53–4) writes that in this language adjectival notions are normally expressed in the form of ‘a separate predication’ and that ‘adjectival verbs may incorporate the nouns they modify to form a single constituent. The resulting complex verb can then function either as a predication or as a nominal’: Cayuga (Mithun 1992: 54)

- | | |
|---|---|
| (1) Ne: ¹ kV he ² hne: ³ wakyes 'aké ne ⁴ a:sató:wa:t | it.is just also CONTR it.is.easy the you.would.hunt |
| ‘Also, it’s an easy way to hunt, as well’ | |
| <hr/> | |
| (2) akya 'tawi'thr-í:yo: (it-)dress-nice | ‘a/the nice dress’ or ‘a/the dress is nice’ |

A little further (*ibid.* 55) she adds that ‘on the extremely rare occasions when separate constituents appear comparable to adjectives or relative clauses in other languages, the modifiers and heads can appear in either order’.

(39) The attributive particle (AP) à occurs with all modifier categories and is ‘morphologically fused onto numerals and the relative particle àmè’ (Noonan 1992: 154). There are, however, also certain differences between verbs and ‘adjectives’ (Noonan 1992: 104f.; personal communication), so that it is perhaps better to treat ‘adjectives’ as a special subset of verbal predicates.

(40) On the verbal character of predicates expressing adjectival notions in Mandarin Chinese, see e.g. Chao (1961: 52) and Gao (1994: 494).

(41) Cf. also Okell (1969: 49–50) and Wheatley (1987: 851) on Burmese, and Saul and Freiberger Wilson (1980: 33) on Nung.

(42) Lee (1989: 40) writes that Korean has a ‘very small’ class of words which can immediately modify the noun. Apparently this includes words expressing adjectival notions (cf. *se ca* [new car] ‘a new car’). If we are indeed dealing with adjectives there cannot be more than one or two of them, since the ‘very small’ class they are part of also includes ‘deictic adjectives’, ‘numeral adjectives’, and ‘interrogative adjectives’ (*ibid.* 59). Compare also Yoo (1997) as well as Lee (1993), who regards -(ü)n and -(ü)l as the realis and the irrealis attributive marker, respectively.

(43) Holton (1999: 342) argues that in Tobelo (which is closely related to Galela) ‘lexical categorization of property words ... depends on the pragmatic status of the nouns they modify. That is, a property word is coded as a noun when the noun it modifies represents “new” information, and as a verb when the noun it modifies represents “old” information’ (cf. also Wetzer 1996 on so-called switch-adjective languages).

(44) On the difference between the two third person forms *i* and *da*, see van Baarda (1908: 60, 84 f.).

(45) 'V is glossed as (1) any vowel, (2) a vowel accented or nasalized by an infix or following suffix (Kimball 1991: p. xxiv). On Koasati phrase terminal markers (PTM), see Kimball (1991: 221–5).

(46) On the expression of adjectival notions in Nivkh, see also Jakobson (1971a: 79) and Beffa (1982: 67, 69, 74).

(47) Note furthermore that in Tsou the present participle form of a verb does not appear to be different from the unmarked (active) verb form (Szakos 1994: 136). Nevertheless, both Elizabeth Zeitoun (personal communication) and Stanley Starosta (personal communication) also state that Tsou does not have a distinct class of adjectives and that the equivalent of English adjectives is simply (nominalized stative) verbs.

(48) 'Doubling' means that members of a certain modifier category are found on both sides of the head constituent (Hawkins 1983: 213).

(49) This study confirms the hypothesis put forward in Hengeveld et al. (forthcoming), which says that doubling of the adjective is dispreferred in languages in which the adjectives do not constitute a clearly differentiated word class (types 1 and 2; cf. Chapter 1). On the difference between AN and NA order, see e.g. Tesnière (1959: 19–23); Bolinger (1967, 1972); Lehmann (1982b: 263); Haiman (1988: 374–5).

(50) 'In general, the Ngalakan NP exhibits a fairly loose sort of structure. It is possible for constituents of what could be considered the "same" NP to be separated from each other by other clausal constituents, or for many NP constituents having the same referent to be strung together in a fairly loose sort of appositional structure' (Merlan 1983: 83).

(51) Nivkh is a rather special case since it only uses classifiers with the first five numerals (section 5.2.2.1.2).

5 Quantifying Modifiers in the Noun Phrase

5.1. Introduction

This chapter is concerned with constituents of the NP that specify quantitative properties of the referent: number markers and numerals. I will mostly deal with plural markers and not so much with finer distinctions in this domain such as dual, trial, quadral, or paucal number (if these are indeed manifestations of number; see section 5.2.1.2). Cardinal numerals are often grammatical elements but there are also quite a few languages in which cardinality is expressed by means of a construction that involves a lexical element, like a verb or a noun. Section 5.2 discusses the *grammatical* expression of the notion ‘quantity’ in the NP (*quantity operators*), section 5.3 the ways in which this notion is given *lexical* expression (*quantity satellites*).

5.2. Number and cardinality: the grammatical expression of the notion Quantity in the noun phrase

5.2.1. Number

I argued in Chapter 2 that a noun does not just designate a property in the spatial dimension, but that its lexical meaning also specifies a particular *Seinsart* (lex-icalized nominal aspectuality), i.e. how that property is represented in the spatial dimension in terms of the aspectual features Shape and Homogeneity. Six *Seinsarten* were distinguished (general, sort, mass, singular object, set, and collective) but two of these are not used in connection with single discrete objects in the physical world: mass and collective. Of the remaining four nominal subcategories only one is marked for number: singular object nouns. The other three are transnumeral: general nouns, sort nouns, set nouns (recall that general nouns are not attested in the current sample). Compare:

general noun (Yucatec Maya; Lucy 1992: 74):

(1) a. *há'as*

banana entity; e.g. banana fruit(s), banana tree(s), banana leaf or leaves, bunch of bananas, banana ‘stuff’ (i.e. the mass), etc.

(p.147)

(1) b. 'un-p'éel	<i>há'as</i>
‘one-CLF:3-dimensional banana(s)’	
e.g. ‘a banana’ (i.e. the fruit)	

sort noun (Mandarin Chinese; Iljic 1994: 100):

(2) a. <i>shū</i>	b. <i>wǔ běn shū</i>
-------------------	----------------------

book(s)	five CLF book(s)
---------	------------------

'book, books' 'five books'

set noun (Oromo; Stroomer 1987: 107):

(3) a. gaala	b. gaala lamaani
camel(s)	camel(s)two
'camel, camels'	'two camels'

singular object noun (Dutch):

(4) a. stoel	b. drie	stoel-en
chair	three chair-PL	
'(one) chair'	'three chairs'	

Since sort nouns and general nouns designate properties that are not characterized as being spatially bounded (e.g. *shū* in the example above might be glossed as 'bookness' in pseudo-English), such nouns cannot be marked for number nor can they enter into a direct construction with a cardinal numeral. Instead a special construction is used, the classifier phrase, in which the numeral is in construction with a classifier (sections 2.2.3 and 2.3.1.3).

The property designated by a set noun (such as Oromo *gaala* 'camel(s)' in the example above) is characterized as being spatially bounded (+ Shape), but it is ambiguous with respect to the feature Homogeneity (cumulative/agglomerative and divisible): although it is always possible to increase the size of a set by adding one or more members to a (singleton or collective) set, only a collective set is divisible. The exact size (number of individuals) of a set may be specified by a cardinal numeral, but since the numeral does not multiply the set entity, we do not get more sets; therefore there is no plural marking. Recall that set nouns can be disambiguated by a nominal aspect marker to specify whether reference is made to a singular object (singleton set) or a collective (collective set).

Lastly, a singular object noun (such as Dutch *(een) stoel* '(one) chair') designates a spatially bounded and indivisible property of a singular entity. A cardinal numeral ($n \geq 2$) or a plural operator will multiply that bounded indivisible space, so that we get more entities with that particular property (Chapter 2). This is expressed through a compulsory plural marker (*twee stoel-en* [two chair-PL] 'two chairs').

Table 5.1 indicates for each language in the sample whether or not nouns are marked for number by specifying which nominal subcategories are used in connection with a singular object in the physical world (consequently there is (p.148) (p.149)

Table 5.1. Nominal subcategories and reduplication

Language	Sg.object noun	[REDUPL]	Set noun	Sort noun
----------	----------------	----------	----------	-----------

Abkhaz	+	?	
<hr/>			
Alamblak (section 2.2)			
<hr/>			
Babungo	+		
<hr/>			
Bambara (section 2.2)			
<hr/>			
Basque		+	
<hr/>			
Berbice Dutch Creole (section 4.2.1.2)	+	+	
<hr/>			
Bukiyip	+		
<hr/>			
Burmese			+
<hr/>			
Burushaski	+	+	+
<hr/>			
Cayuga			
<hr/>			
(see section 1.5.2)			
<hr/>			
Chinese, Mandarin (section 5.2.1.3)		? +	
<hr/>			
Chukchi		+	
<hr/>			
Dutch	+		
<hr/>			
Galela		+	
<hr/>			
Georgian		+	
<hr/>			
Guaraní (section 2.2)			
<hr/>			
Gude		+	
<hr/>			
Hittite	+	+	
<hr/>			
Hixkaryana		+	
<hr/>			
Hmong Njua (section 4.2.1.2)		+	
<hr/>			
Hungarian		+	
<hr/>			
Hurrian (section 2.2.4)			
<hr/>			
Ika			
<hr/>			
Kayardild	+	+	
<hr/>			
Ket (sections 2.2.1 and 4.2.1.2) ¹	+		?
<hr/>			
Kisi	+		
<hr/>			
Koasati (section 2.2)			
<hr/>			
Korean			+
<hr/>			
Krongo	+	+	
<hr/>			
Lango		+	
<hr/>			
Nahali		+	
<hr/>			
Nama (section 2.2)			
<hr/>			
Nasioi (section 2.2.1)	+	+	
<hr/>			
Ngalakan		+	
<hr/>			
Ngiti	+	+	
<hr/>			

Nivkh (sections 2.2.2 and 5.2.2.1.2)		+	?
Nung			+
Nunggubuyu (section 4.2.1.2)	+ ?	+	+
Oromo			+
Pipil (section 2.2.1)	+ ?	+	?
Quechua, Imbabura (see section 2.2.4)			
Samoan (see section 2.2.4)		+	
Sarcee (sections 2.2.1 and 4.2.1.2)	+ ?		+
Sumerian		+	+
Tamil	+		+
Tsou		+	+
Turkish (sections 2.2.1 and 4.2.1.2)			
Vietnamese			+
Wambon	+	+	
West Greenlandic	+		

(1) On the so-called singulative affixes in Ket, see Werner (1997a: 99).

considerable overlap with the information contained in Table 2.1). Recall that strictly speaking number marking is only possible with singular object nouns (and collective nouns—which are ignored here). Since some language use reduplication (REDUPL) to mark number this is also included in Table 5.1 (but see section 5.2.1.2).

5.2.1.1. Languages with number marking

For the following languages in the sample it could be established with reasonable certainty that they have a (small or large) group of nouns which must be marked for number, both with and without a modifying numeral: Abkhaz, Babungo, Bukiyp, Burushaski, Dutch, Hittite, Ket, Kisi, Krongo, Nasioi, Ngiti, Tamil, and West Greenlandic.¹ In the following languages singular number receives no formal expression (as in English), i.e. the unmarked form is inherently characterized for singular number: Abkhaz, Burushaski, Dutch, Hittite, Ket, Nasioi,

(p.150) Tamil, and West Greenlandic. Here are some examples from Abkhaz (on the article, see sections 3.3.2.4.1 and 6.2.1.3) and Burushaski:

Abkhaz (Hewitt 1979: 57, 237)

(5) a. à-jyab	b. à-jyab-c°a	pš-y°o(-k')
ART-girl	ART-girl-PL	four-HUM.AX-(one)
'(the) girl'	'(the) four girls'	

Burushaski (Lorimer 1935–8: 1. 15, 187)

(6) a. hin	hir	b. altán	híri
------------	-----	----------	------

one man	two men
'a man'	'the two men'

In the other languages both singular and plural number are overtly expressed. In Babungo, Bukiyp, Kisi, Nasioi, and Ngiti number marking is closely related to class membership.²

Kisi (Childs 1995: 14)

(7) a. nì-leŋ	b. nì-láŋ
ear-CL.SG	ear-CL.PL
'ear'	'ears'

Most nouns have a double-class gender in that they belong to one class in the singular (say, class 3 or CL3) and another (class 4 or CL4) in the plural, but in some cases members of a class remain unmarked, as in these examples from Babungo (see also examples from Sesotho given earlier in section 3.3.2.2.2):

Babungo (Schaub 1985: 13–4)

(8) a. φ-lwàŋ	mù'	b. yí-lwáŋ	tēe
CL3-hammer	one	CL4-hammer	five
'one hammer'		'five hammers'	

As in the three Niger-Congo languages, nouns in Bukiyp also occur with a noun class affix which simultaneously marks singular or plural number:

Bukiyp (Conrad 1991: 10)

(9) a. nyba-t	b. nyba-gu
dog-CL11.SG	dog-CL11.PL
'dog'	'dogs'

In the case of Babungo, Bukiyp, and all other languages that obligatorily mark singular and plural number (both with and without a modifying numeral) one could ask whether we are dealing here with singular object nouns (i.e. true number marking) or erstwhile set nouns with lexicalized singulative and collective aspect (p.151) marking. This is not an easy question to answer, but since the unmarked (class/numberless) form does not seem to occur by itself as the head of an NP in actual speech, I have assumed that in the case of Bukiyp and the Niger-Congo languages (Babungo, Bambara, Kisi) both the singular and the plural form of the noun are stored as such in the lexicon (constituting a so-called 'regularity' in the lexicon (see Dik 1997: 342–5; cf. also Anderson 1985: 174f.).

In Ngiti most nouns are set nouns, but 'nouns denoting humans, and any compound form whose second part is in fact originally a [+human] noun ... and any compounds

with the inalienable nouns *-akpà* “male”, *-àyi* “female humans” [etc.—JR]’ are marked for number. Ngiti has three strategies of plural marking: suppletion, the use of Bantu-like prefixes *mU-* and *pba-*, and tone (Kutsch Lojenga 1994: 133 f.):

Ngiti: suppletion (Kutsch Lojenga 1994: 133)

- (10) a. *ingba* b. *inzo*
‘child’ ‘children’

Plural marking through (high) tone is restricted to a small group of human nouns (predominantly loan words; e.g. *màlimò* derives from Swahili *mwalimu*). Note that in (11b) the last two syllables are replaced by high tones:

Ngiti: tone (Kutsch Lojenga 1994: 135)

- (11) a. *màlimò* b. *màlímó*
‘teacher’ ‘teachers’

The prefixation strategy is most probably due to the fact that speakers of Ngiti are familiar with certain Bantu languages so that the prefix *mU-* reminds them of the Bantu class 1 prefix, which typically occurs with human nouns. Consequently nouns ‘denoting human beings are sometimes found with the prefix *pba-*, replacing *mU-*, to indicate the plural’ (*ibid.* 134):

Ngiti: prefixation (Kutsch Lojenga 1994: 134)

- (12) a. *mùtànì* b. *pbàtàhì*
‘cowherd’ ‘cowherds’

In addition to a very large group of set nouns Krongo also has a small group of nouns which only occur with overt singular and plural marking (Reh 1985: 102; cf. also note 1 on lexicalized forms containing an erstwhile nominal aspect marker in Oromo):

Bei einer weiteren Gruppe von Nomina, die vor allem Bezeichnungen für Werkzeuge und Gefäße umfaßt, wird die morphologisch unmarkierte Form entweder überhaupt nicht oder nur in augmentativer Bedeutung gebraucht. Singular- und Pluralform sind in diesem Fall beide offen markiert.

(p.152)

[With another group of nouns, notably those denoting instruments and containers, thi morphologically unmarked form is either not used at all or only in an augmentative sense In either case the singular and plural markers are overtly expressed.]

ECrongo (Reh 1985: 102)

- (13) a. *tìn-kándá* b. *nì-kándá*

SG-spear	PL-spear [my glosses—JR]
'spear'	'spears'

The remaining languages with an entry under Singular Object Noun are all more or less doubtful cases; they were discussed in sections 2.2.1 (Pipil) and 4.2.1.2 (Nunggubuyu, Sarcee).

Let me finally point out that in many languages number marking (if it is that, for in many cases I believe we are actually dealing with nominal aspect marking) is often highly irregular, even to the extent that, for example, native speakers of Oromo are 'often in doubt or in disagreement about the correct form of the plural noun' (Stroomer 1995: 41). The Nilo-Saharan languages are especially notorious in this respect (Tucker and Mpaayei 1955: 4).

5.2.1.2. Reduplication

At least the following languages in the sample use reduplication to express plural number: Berbice Dutch Creole (Kouwenberg 1991: 171), Burushaski (Berger 1998: 223), Ika (Frank 1990: 56), Kayardild (Evans 1995: 200), Ngalakan (see below), Nunggubuyu (Heath 1984: 197), Pipil (Campbell 1985: 51), Samoan (Mosel and Hovdhaugen 1992: 218 f.), Sumerian (Thomsen 1984: 59–60), Tsou (Tung 1964: 169; Szakos 1994: 44 f.), and Wambon (see below).³

It is not always the noun that is reduplicated to express the notion of plurality; in some languages it (also) concerns the adjective or the modifying verb (as in e.g. Ika, Pipil, Wambon, Boumaa Fijian).⁴ In the case of Sumerian, however, the meaning of reduplication (which only occurs with inanimate nouns) was rather to express 'totality'. Reduplication is often restricted to certain groups of nouns. In Ngalakan, for example, reduplication only occurs with kin terms (Merlan 1983: 53) and Wambon has only one noun that can be reduplicated: *kap* 'man, person' (de Vries 1989: 36). It seems that reduplication is predominantly a feature of set nouns (sections 2.2.2 and 4.2.1.2), which in its turn would suggest that reduplication is perhaps a way to mark nominal aspect rather than number.

The same may hold for dual marking, which also seems to be restricted to languages with set nouns (or undecided cases such as West Greenlandic). In the (p.153) sample, dual marking in the NP (i.e. not on pronouns or in the verbal complex) is attested in at least the following languages: Nahali, Nama Hottentot, Nasioi, Nunggubuyu, West Greenlandic (northern dialects only). There are certain arguments in favor of the view that the dual should be interpreted as a special kind of collective aspect marker.⁵ For instance, in many languages the dual suffix is clearly diachronically related to the equivalent of the numeral 'two' (Dixon 1980: 323; see also Donaldson 1980: 102, 264, who regards the dual marker as one of the 'collective suffixes'), and it occurs with the unmarked form of the noun, i.e. there is no sign of an additional plural or collective aspect marker. Finally, there are several examples of languages in which dual marked NPs (just like NPs headed by set nouns; cf. section 4.2.1.2) are cross-referenced as 'singular' in the verbal complex (see also e.g. Whorf 1946: 175 on Hopi).⁶

Mangarayi (Merlan 1982: 91)

(14) φ-!ur?+ma-ñ	wađij	ŋabarana	jadba-ra-ŋan
3SG/3SG-butcher-PAST.PRT	also	two	upper leg-DU-ACC
'He also butchered/cut up the upper two legs'			

Instances of so-called number marking on set nouns have already been discussed in Chapter 3, so that we can now turn directly to the last column, which concerns languages that employ sort nouns.

5.2.1.3. Number and sort nouns

Sort nouns are attested in the following languages (section 2.2.3): Burmese, Mandarin Chinese, Korean, Nung, and Vietnamese.⁷ Although sort nouns cannot be marked for number, we do find statements in the grammars of these languages claiming that number marking occurs.⁸ Closer inspection reveals, however, that the so-called plural markers in these languages have little in common with the obligatory markers of nominal plurality as they are found in e.g. Dutch and English (cf. Austerlitz 1980: 242). I will argue below that the so-called plural markers are better regarded as quantifiers (equivalent to English quantifiers such as ‘several’, ‘some’, and ‘many’) and that in Mandarin Chinese (and possibly Korean) we are actually dealing with a collective (aspect) marker rather than a plural (number) marker.

The element that is often regarded as a plural marker in Mandarin Chinese is the suffix *-men*, which also implies definiteness. However, it is used only (p.154) occasionally and its occurrence is restricted to human nouns (cf. Norman 1988: 159):⁹

Number is obligatorily expressed only for the pronouns. The same plural suffix found in the pronouns, *-men*, can also be employed with nouns referring to human beings; however, the resulting forms differ from English plural nouns in several ways. They are not used with numerals; they are not obligatory in any context; and they tend to refer to groups of people taken collectively. Examples: *háizimen* ‘(a certain) group of children’, *lǎoshīmen* ‘the teachers’.

These remarks indicate that *-men* marks collectivity rather than plurality, which suggests that human nouns have changed their subcategorial status (or are in the process of doing so) and should be regarded as set nouns.

Several studies have been devoted to the rare and ‘versatile’ Korean plural marker *tul*, but many of its syntactic and semantic properties remain unclear (cf. Seok Choong Song 1975; Kuh 1986; Lee 1991; Kim 1993; Kang 1994; Sohn 1994: 268–9). For example, it can occur in many different places in the sentence and although it is clear that *tul* cannot be characterized as a plural marker (Song 1997), there seems to be no consensus as to its proper meaning(s). Some state that it means ‘all, several, together’ (Ramstedt 1939, as quoted in Seok Choong Song 1975: 538) or ‘group’ (Martin and Lee 1969: 32), whereas others have argued it has a distributive meaning or serves as a focal element (Song 1997).

Below are some examples of the so-called plural marker in Burmese, Nung, and Vietnamese. They are optional and do not seem to convey the notion of ‘group’ or ‘collective’. Furthermore the way they are characterized in the various grammatical descriptions (see below) suggests that they are better treated as members of a closed class of quantifiers (like English ‘some, few’; cf. also Schachter 1985: 38).

Burmese (Okell 1969: 96)

(15) eiñ	hka?	cicì	tei
house	fairly_big	PL	
'fairly houses'			

Nung (Saul and Freiberger Wilson 1980: 29)

(16) mahn	má	hau	lái
PL	dog	bark	much
'the dogs are barking quite a lot'			

Vietnamese (Emeneau 1951: 87)

(17) nhū'ng	con	trâu
PL	CLF	buffalo
'the buffaloes'		

(p.155) In Burmese, where they are called ‘auxiliary nouns’ (Okell 1969: 82), the three ‘plural markers’ *toú*, *tei/twei*, and *myà* (the differences are not made explicit) are discussed alongside such apparently varied items as *taiñ* ‘every’, *cì* ‘great’, *hkalei/lei* ‘little’, *hsouñ* ‘extreme’, *lau?* ‘approximately’, and *htè* ‘only, no more’.¹⁰ In Nung *mahn* ‘plural’ is one of the ‘non-specific numerators’ (Saul and Freiberger Wilson 1980: 22); some other members of this class are *i* ‘small amount/little’, *kì* ‘several’, *lái* ‘much/many’, *táhc* ‘any’, and *thém* ‘more’. The Vietnamese marker for plurality *nhū'ng* is called ‘general quantifier’ (*ein all-gemeiner Quantifikator*—Kölver 1982: 170) and ‘pluralizer article’ (Nguyễn Đăng Liêm 1969: 113; cf. also Thompson 1965: 179; Emeneau 1951: 87–8) and falls in the same category as, for instance, *các* ‘plural generality’, *mọi* ‘every’, and *máy* ‘restrictive plurality, some’.¹¹

In sum, number marking as we find it with singular object nouns is not quite the same as instances of so-called number marking on set nouns and sort nouns. In the case of set nouns we are actually dealing with nominal aspect markers, which specify what kind of set the speaker is referring to (a singleton set or a collective set). In the case of sort nouns we are dealing with members of the category of quantifiers.

5.2.1.4. Number and nominal aspect marking

Number marking is still largely *terra incognita*, especially from a cross-linguistic point of view. As long as sufficient and reliable information is generally lacking for most languages, it will be rather difficult to come to grips with the intricacies of number marking in general (cf. Corbett 2000). Nevertheless, the distinction between number and nominal aspect marking may explain certain properties of (real and apparent) number markers and some other problems as well, such as ‘number disagreement’ and semantic properties of incorporated and predicate nouns (Chapter 4).

(p.156) Some of the problems encountered in connection with nominal aspect and number marking are due to the fact that languages are dynamic systems. It will often be the case that one or more parts of the system are undergoing some sort of change. This means, of course, that nouns can be in transition as regards their aspectual status (see e.g. Oromo and Mandarin above). As nouns may be in the process of changing their *Seinsart*, we will inevitably come across languages with a group of nouns that are in an intermediate phase, displaying characteristics of different aspectual types. This can make it difficult to draw a hard and fast line between number and aspect marking. In fact a similar situation holds true as regards tense and verbal aspect. Despite the fact that tense and verbal aspect have been thoroughly investigated in many different languages (which cannot be said about number marking, not to mention nominal aspect), there are still many problems that have not been solved, some of them very fundamental. For instance, it is still not entirely clear whether the perfect is a tense or an aspect (Comrie 1976: 5–6, Chapter 3; 1985: 33).

5.2.2. Cardinality

The notion Quantity can be expressed by means of number distinctions, but also in the form of cardinal numerals. This section discusses some properties of grammatical expressions of cardinality; quantifiers (such as ‘many’, ‘few’, ‘all’, ‘some’) are ignored in this study, mainly because sufficient and reliable data are scarce.

Many, perhaps most languages have numeral systems, i.e. linguistic expressions for cardinal numbers. Some such languages only use simple expressions for numbers (like ‘one’, ‘six’, or ‘eleven’; also called the *atoms* of a numeral system) and do not make use of arithmetic operations (such as addition as in e.g. ‘seventeen’; $7 + 10$); the highest value in such simple systems is reported to be 5 and the lowest 2 (Greenberg 1978b: 256). In the numeral systems of other languages atoms are used to construe complex expressions for higher numbers. According to Greenberg (1978b: 257–8) all four fundamental arithmetic operations are used for this purpose: addition, subtraction, multiplication, and division. Thus, in English 23 is expressed as a function $(a \times b) + c$ in which a has the value 10, b has the value 2, and c has the value 3. In this book, however, I am not so much interested in the internal organization of numeral expressions as in their external morpho-syntactic properties.¹²

There are also languages which, strictly speaking, have no numeral system (Wilkins 1989). Instead they may use expressions that must be accompanied by a gesture (often the gesture alone will suffice). For instance, in Kobon (Indo-Pacific; Davies 1981: 206–7; see also Greenberg 1978b: 257). (p.157)

the measuring code against which counting is done is a sequence of body parts starting with the little finger on the left hand and moving up the arm to the shoulder, across to the other shoulder, and down the other arm, and then back in the reverse direction. The relevant body-part is indicated by touching it with the forefinger of the opposite side. The name of the body-part is sometimes spoken at the same time as it is touched in which case the names of the body-parts are followed by the word *böng* ‘one side of when going through the sequence of the body-parts for the first time, i.e. from one to twenty-three. When going through for the second time, i.e. from twenty-four to forty-six (in reverse direction from right side to left side), the names of the body-parts on the left side of the body are followed by *böng*, and so on for the third and subsequent times.

Thus, *mudun* ‘forearm’ is used to express the numbers 7, 30, 53 etc., and *mudun böng* for 17, 40, 63 etc. It is doubtful whether such expressions are used as modifiers in an integral NP.¹³

Even in languages that are reported to have numerals, the role of body-parts is sometimes easily recognized. For instance, in Cuna, and indeed in many other Amerind languages, 20 is expressed as ‘one man’ (*tulakwena*), referring to a person’s total number of fingers and toes. In a similar vein 40 is ‘two men’ (*tur-pokwa*) and 100 ‘five men’ (*tul-attale*; Holmer 1947: 102). The Australian language Kayardild, which only has numerals up to ‘four’, is also a case in point. According to Evans (1995: 242) ‘counting is not a traditional activity in Kayardild society; in fact, the counting of turtle eggs was, and is still, tabooed. The only time I heard a higher number used was in a discussion of how many wives a man had; the expression *kiyarrngka marlda* “two hands” was used for “ten.”’

Table 5.2 summarizes various kinds of information in relation to numeral expressions, involving both cardinal (num) and ordinal (ord) numerals, which is relevant in connection with the principles of constituent order, to be discussed in Chapters 8–10. The first column gives the position(s) of the numeral relative to the head noun (N). There may be several possibilities here. For instance, in certain languages all cardinal numerals either precede or follow the head noun; in other languages cardinals may occur on either side of the noun, and in yet another group of languages some cardinals must precede and others must follow the head noun. Such details are not indicated in Table 5.2, but are discussed in section 5.2.2.1.2. Notice that in some languages cardinal numerals co-occur with a numeral classifier (NUM CLF); these languages will be treated more elaborately in section 5.2.2.1.3.

If cardinal numerals do not (all) constitute grammatical elements, this is indicated in the second column (V = verb, N = noun, adv = adverb; P means that numerals must sometimes or always be used in a predicative construction).

The position of the ordinal numeral relative to the head noun, if known, is given in the third column, showing that in most cases its position coincides with that of the cardinal. The last column indicates how ordinals are productively (p.158)

Table 5.2. Noun-numeral constructions

Language	Numeral-N order	Category numeral	Ordinal-N order	Ordinal formation
Abkhaz	num(-)N num		ord N	suffix/id
Alamblak	num N num		ord N ord	
Babungo	N num	N	ord N	id
Bambara	N num	(N)	N ord	suffix
Basque	num N num		ord N	suffix
Berbice Dutch Creole	num N			
Bukiyip	num N num			
Burmese	N num CLF			
Burushaski	num N		ord N	suffix

Cayuga				
Chinese, Mandarin	num CLF N		ord N	prefix
Chukchi	num(-) N	V	ord N	suffix
Dutch	num N		ord N	suffix
Galela	N CLF? num	P	N ord	ma num
Georgian	num N		ord N	circumfix
Guaraní	num N			
Gude	N num		ord N	suffix nga
Hittite	num N			
Hixkaryana (section 5.3.2)	num N	adv		
Hmong Njua	num CLF N			
Hungarian	num N			
Hurrian		N		suffix
Ika	num N num			
Kayardild	num N			
Ket	num N		ord N	suffix
Kisi	N num			
Koasati (sections 2.2 and 5.3.2)		V		prefix/id
Korean	num (CLF i) N num CLF			suffix/prefix
Krongo	N num	V		id
Lango	N num			
Nahali	num N (ex)			
Nama Hottentot	num N		ord N	suffix
Nasioi	N num			
Ngalakan	num N			
Ngiti	num N num		ord N (but see below)	
Nivkh	num(:CLF)-N- num:CLF		(see below)	
Nung	num CLF N num		Nord	tai num
Nunggubuyu	num N num	N		
Oromo	N num	N?	Nord	k/taa num
Pipil	num N		ordN	
Quechua, Imbabura	num N		ordN	suffix
Samoan (section 5.3.1)	N e num	P	Nord	
Sarcee	num N num	V?		

Sumerian	N num	P	Nord	GEN + COP
Tamil	num N num		ordN	suffix
Tsou	num N			
Turkish	num N		ordN	suffix
Vietnamese	num CLF N num CLF			
Wambon	num-o N num			
West Greenlandic	N num	N	Nord	

(p.159) formed from cardinals (sx = suffix, px = prefix). Sometimes, however, a special construction or an extra element (e.g. Galela *ma*) is employed and some languages do not (always) appear to distinguish formally between cardinals and ordinals at all (= id). In the last two columns a blank indicates that no conclusive information could be obtained and (ex) means that the information is based on examples, not on explicit statements; question marks are explained below (Galela section 5.2.2.1.3; Sarcee section 5.3.1; Hurrian and Oromo section 5.3.2).

5.2.2.1. Formal properties of cardinal numerals

5.2.2.1.1. Bound versus free cardinals

In most languages numerals are free forms, but there are some languages in which the numeral may or must be attached to some other constituent, usually the noun. In the sample this happens in the following languages: Abkhaz, Basque, Chukchi, Hurrian, and Nivkh (sections 5.2.2.1.1 and 5.2.2.1.3).

In Abkhaz the numeral is only fixed directly to the root in the case of a (following) non-human noun (Hewitt 1979: 236 f.; Spruit 1986: 51; Siewierska et al. 1997: 788, 797):

Abkhaz (Hewitt 1979: 236–7)

(18) a. pš-y°ə(-k')	à-jyab-c°a	b. à-jyab-c°a	pš-y°ə(-k')
four-HUM(-one)	ART-girl-PL	ART-girl-PL	four-HUM(-one)
'(the) four girls'		'(the) four girls'	

(p.160)

(19) a. pš-la-k'	b. a-la-k°a	pš-bà
four-dog-one	ART-dog-PL	four-BÀ
'(the) four dogs'		'(the) four dogs'

In Basque ‘one’ and (in certain dialects) ‘two’ are phonologically part of the head noun (Saltarelli 1988: 78) and in Hurrian ‘one’ can be expressed as a bound form (Speiser 1941: 82). In Chukchi numerals must form a compound with the head noun if the NP does not occur in the absolute case (Vladimir Nedjalkov, personal communication; see also Kämpfe and Volodin 1995: 102).

Chukchi (Vladimir Nedjalkov, personal communication)

(20) etlon	ga-twetcha-twa-len	ga-ŋəron-elg-ə-qaa-ma
he	PERF-stand_up-be-3SG	COM-three-white-ə-deer-COM

'He stood next to the three white deer'

5.2.2.1.2. Doubling of numerals

The sample contains various languages in which a cardinal numeral may precede or follow the head noun: Abkhaz, Alambalak, Basque, Bukiyp (section 4.3.4), Nivkh, Ika, possibly Kayardild (see 4.3.4), Korean, Ngiti, Nung, Nunggubuyu, Sarcee, Tamil, Vietnamese, and Wambon.¹⁴ Doubling in classifier languages is discussed separately in section 5.2.2.1.3.

In the case of Alambalak, Nunggubuyu, Sarcee, Vietnamese, and Wambon it is not clear which are the factors that determine the actual position of the cardinal.¹⁵ Regarding the position of adnominal numerals in Wambon, for example, de Vries (1989: 43) writes:

The numerals may occur both pre- and postnominally; we have not found semantic or pragmatic differences associated with pre- and postnominal occurrence. In prenominal occurrence we sometimes find *takhemo* 'three' instead of the longer form *ilumtakhemo* ... When the numerals precede the noun, they take the connective *-o* just like all other prenominal modifiers.

Wambon (de Vries 1989: 42–3; TN = transitional nasal)

(21) a. ap	ndominuk
house	one
b. ndominukh-o (-n-)	ap
one-O	(TN)

'one house'

(p.161) As for ordering preferences, in Alambalak (Bruce 1984: 99), Sarcee (Cook 1984: 76), and Vietnamese (Nguyễn Đình Hoà 1987: 785) the numeral most commonly precedes the noun.

In Abkhaz the position of the cardinal numeral is influenced by the class of the head noun. Whereas numerals do not seem to have a positional preference in the case of a human noun, they tend to follow non-human nouns (see above). The numeral 'one', however, which is formally identical to the indefiniteness marker, always seems to appear as a suffix (Hewitt 1979: 153, 236 f.).

In Basque only the numeral 'one' appears after the head noun, although in certain dialects the equivalent of 'two' *bi* also follows.¹⁶ All Nung numerals (plus classifier) precede the noun, except most equivalents of 'one', which may assume different forms. One of these, however, precedes the noun (Saul and Freiberger Wilson 1980: 14, 21–2).

As stated earlier (note 7), only numerals up to ‘five’ require a classifier in Nivkh. Formerly the numeral classifier phrase only used to follow the head noun, but according to Panfilov (1962–5: i. 191) there is now a tendency for the numeral classifier phrase to take up the prenominal position used by the higher numerals, which appear without a numeral classifier (Comrie 1981a: 269). The original examples are due to Panfilov (1962–5: i. 192) and Krejnovič (1934: 204), respectively.

Nivkh (Mattissen and Drossard 1998: 51–2; see also Gruzdeva 1998: 25)

(22) mu-ť'om-γir	ťo-ŋəŋ-d-γu
boat-five:CLF-INS fish-hunt-IND/NML-PL	
'They went fishing in five boats'	

(23) ŋax-kumusk
six-ruble
'six rubles'

The position of numerals in Ika is determined by the pragmatic status of the referent of the NP: numerals precede the head noun in indefinite NPs and occur after the noun in NPs with definite reference (Frank 1990: 31).

The unmarked position of the numeral in Ngiti is before the noun, but both in connected speech and in written texts it is also found in postnominal position. In fact, it can even occur by itself in clause-initial position (see also Chapter 8 on discontinuity). According to Kutsch Lojenga (1994: 354) ‘the different word orders signal different degrees of emphasis. Numeral-noun is the unmarked order with least emphasis, noun-numeral is slightly emphasised, and the numeral in (p.162) clause-initial position is the most emphatic’; compare

Ngiti (Kutsch Lojenga 1994: 355)

(24) ma	m-ጀkጀrጀ	àrጀ	tsu
1SG SC-cut:PERF.PRES	eight	tree	
'I have cut eight trees'			

(25) ma	m-ጀkጀrጀ	itsu	àrጀ
1SG	SC-cut:PERF.PRES	tree	eight
'I have cut eight trees'			

(26) àrጀ	ma	m-ጀkጀrጀ	itsu
eight	1SG	SC-cut:PERF.PRES	tree
'I have cut eight trees'			

Furthermore, there are several numerals that have different tones depending on whether they precede the noun or not. For example, adnominal *ɔyɔ* ‘two’ becomes *ɔ᷑ɔ* (with high tone and low-mid tone, respectively) when it occurs elsewhere in the sentence (*ibid.* 355).

As a rule the numeral is placed after the noun in Sumerian, but ‘for practical reasons’ it often precedes in economic texts (Thomsen 1984: 82–3, 276). However, since the special syntax of numeral-noun constructions in (written) economic texts is normally absent or at most a very marginal phenomenon in spoken language (see Greenberg 1975: 32–3 on listing constructions), I have decided not to include this variant in Table 5.2.

In Tamil numerals normally precede the head noun, but they may also follow, in which case the ‘nominal form’ of the numeral is used (at least ‘in cases where there are different nominal and adjectival forms’ (Asher 1982: 66–7); only the nominal form inflects for case).¹⁷

5.2.2.1.3. Numeral classifier languages

It will be remembered that sort nouns cannot be in a direct construction with numerals, because the properties they designate are not characterized as having a boundary in the spatial dimension (section 2.4.1). For that reason languages with sort nouns employ a special construction to specify the cardinality of a referent: the numeral classifier phrase, in which the numeral is not in a direct construction with the noun, but rather with a so-called numeral (or: sortal) classifier. Numeral classifiers typically occur with numerals and often provide some sort of semantic classification of the head noun. In the sample classifier constructions are attested in Burmese, Hmong Njua (but cf. section 4.3.5), Mandarin Chinese, Korean, Nivkh (but see above), Nung, and Vietnamese (see also section 2.2.3). In at least two of these languages, namely Burmese (Wheatley 1987: 851) and Korean (Lee 1989: 117–18), the (post-nominal) classifier phrase is deemed to be an apposition (cf. also note 23 in Chapter 1).

(p.163) Galela constitutes a rather special case for two reasons. First, the classifying element is rarely used (which is why I have assumed that Galela employs set nouns rather than sort nouns). Secondly, it precedes the numeral, which is very unusual when the classifier is a free form (cf. Lehman 1990: 104–5 n. 6). The same ordering pattern is reported for a few other languages, but it is not certain that this always involves a numeral classifier. In at least one such language (Kilivila; Senft 1986) the numeral classifier is probably better categorized as a class marker (Rijkhoff 1990a: 33). The interesting thing about Galela classifiers is that they, too, are preceded by the ‘article’ *o* (see also 6.2.1.3 where it is suggested that it is perhaps better analyzed as a noun marker). For example:

Galela (van Baarda 1908: 44)

(27) <i>o</i>	<i>gunangé</i>	<i>o</i>	<i>làko</i>	<i>iha</i>
ART	sago-cookie	ART	CLF:eye	four

‘four sago cookies’ [these are baked in a device with six partitions; a partition is called a *làko*, i.e. the eyes of the baking device]

Although no definite explanation can be offered on the basis of the data contained in the grammar, it may be the case that together the so-called classifier + numeral constitute another term (notice that in such a case the order would be: (pro)-noun + numeral), which has an appositional relationship with the preceding noun (or rather NP).

The numeral classifier also appears before the numeral in e.g. Cuna (Amerind, Chibchan) but in this language the classifier is always a bound form (e.g. *us warpo* [machango CLF-two] ‘two machangos’, *sappi sar-po* [tree CLF-two] ‘two trees’—Holmer 1947: 101; cf. also Greenberg 1989). Furthermore, the status and function of classifying elements in the Amerind languages seems to be rather different as compared to such elements in (South-East) Asia and the Pacific (cf. Payne 1987; Derbyshire and Payne 1990). For instance, it has been argued that in the classifying systems of Jacalteco and Yagua it is difficult to distinguish between numeral classification and noun class marking (Craig 1986a) and that the classificatory elements are both inflectional and derivational in nature (Payne 1986). This is not to say that in languages spoken outside the American continent the situation is always very clear when it comes to deciding between numeral classifiers and noun class or gender markers (see section 3.3.2). The classification system of Nasioi, for example, has been recognized as an intermediate type (Foley 1986: 84–5), because of its large number of classifying elements (which is rather typical of numeral classifier systems) and the fact that these elements occur with various other constituents in the NP (resembling agreement in gender or noun class systems).

In certain languages the expression of the numeral changes when human nouns are involved. Although in a strict sense this may not be regarded as an instance of numeral classification, it is of course rather peculiar that the class of the (p.164) head noun only affects the expression of numerals. Thus, Sarcee numerals (except *tlik'azá* ‘one’) must be provided with the suffix *-na* in the case of humans (Cook 1984: 76–80). In Abkhaz, Burushaski, and Samoan, the + human distinction in nouns is also reflected in different forms of expression of the numeral (Hewitt 1979: 236; Lorimer 1935–8: i. 178; Mosel and Hovdhaugen 1992: 318). Alongside the common set of numerals Hungarian has another set which is optionally used for persons. This second set has the suffix *-an* or *-en* (the exact form is determined by vowel harmony considerations; section 3.3.2.1) and usually follows rather than precedes the noun and does not seem to be part of the integral NP (Greenberg 1975: 39–40).

Doubling in numeral classifier languages. In most numeral classifier languages in the sample the classifier phrase may only appear either before or after the head noun (Burmese, Hmong Njua, Mandarin Chinese, Nung), but in Korean, Nivkh, and Vietnamese it may precede or follow.¹⁸

In Korean the classifier phrase most frequently follows the noun, but may also precede. In postnominal position it is characterized as a ‘special type of appositional construction’ (Lee 1989: 117–18); if it precedes the noun it must occur with the element *i* and is analyzed as being part of an endocentric construction (see also section 1.5.4).¹⁹

Korean (Lee 1989: 117–18)

(28) ceg isibo gwʌn

book twenty-five volume

'twenty-five books'

(29) isibo gwʌn-ii cəg

twenty-five volume-II book

'twenty-five books'

In Vietnamese the classifier phrase normally precedes the head noun, but apparently the reverse order is also possible (although it is not clear under what circumstances; cf. Nguyễn Đình Hoà 1987: 785).

As noted before (section 4.3.3), Nivkh presents a special case because the NP appears to constitute a polysynthetic construction in which only lower numerals occur with a classifier. Comrie (1981a: 269) writes:²⁰

[Classifiers] function most clearly with the numerals ‘one’ to ‘five’, for which Nivkh nouns are divided into 26 classes, with different forms for each class. For numerals up (p.165) to five, where the classifiers are used, the usual traditional order is for the numeral-classifier complex to follow the noun, e.g. *qan mor* ‘two dogs’, literally ‘dog two-animal’; with higher numerals, used without classifiers, the numeral precedes, e.g. *ŋamg nívx* ‘seven people’. The latter reflects the typical Nivkh order with the attribute preceding the head. The classifier construction also, in a sense, reflects this word order internally to the numeral-classifier complex, and was probably originally an appositive construction with the noun and the numeral-classifier complex in apposition. Currently in Nivkh, there is a tendency for the numeral-classifier to be preposed [Panfilov 1962–5: i. 191], presumably reflecting reanalysis as an attribute-head construction.

Properties associated with prenominal placement. The facts from Korean and Nivkh mentioned above confirm Greenberg’s theory about the diachronic development of numeral classifier phrases (Greenberg 1975 and 1989: 110; see also note 12 in Chapter 3). In this theory the classifier phrase starts out as an adverbial construction at clause level, which gradually (through postnominal apposition) becomes an integrated constituent of the NP, usually ending up before the head noun. As it appears, prenominal position of the classifier phrase often entails other properties, some of which indicate that we might no longer be dealing with ‘pure’ sort nouns. Some such properties are:²¹

- the numeral classifier is no longer restricted to numerals; it may also occur with other NP constituents, such as the demonstrative pronoun (e.g. Mandarin Chinese);²²
- some or all numerals may or must occur without the classifier (e.g. Nivkh; see also Jaisser 1987: 176 on Hmong);
- the numeral classifier no longer serves as an anaphoric element but has come to be associated with such notions as definiteness, specificity, and topicality (e.g. Goral 1978: 14 on Vietnamese; Sun 1988 and Li and Thompson 1989: 130 on Mandarin Chinese; Hopper 1986 on Malay). In some such languages the numeral classifier may occur alone (e.g. Hmong Njua, Nung);²³

- there appears a pluralizing or perhaps better collectivizing element (cf. Li and Thompson 1989: 40–1, Chao 1961: 40, and Norman 1988: 159 on (p.166) Mandarin Chinese; cf. also Adams 1989, Bisang 1996, 1999, Jones 1970, and Goral 1978 on numeral classifiers in South-East Asian languages in general).

Another feature of these languages seems to be that the number of numeral classifiers is reduced (and that this perhaps also marks the beginning of the change from sort nouns to set nouns; section 5.2.1.3). Consider these remarks by Greenberg (1972: 32), especially the last sentence (recall that as a rule nominal aspect markers are mutually exclusive with numerals; sections 2.2.2 and 4.2.1.2).²⁴

In languages in which the demonstrative occurs with classifiers there seems invariably to be a single ‘plural’ classifier replacing the ordinary classifier but only in the demonstrative construction, not in the numerical construction. For example in MANDARIN the classifier *běn* required with *shū* ‘book’ with any number, e.g./*běn shū* ‘this book’ *san běn shū* ‘three books’ occurs with the demonstrative also, *chè běn shū* ‘this book’ but only in the singular. For all nouns, the ordinary classifier is replaced by *hsie* ‘some’ to form the plural with demonstratives, *chè hsie shū* ‘these books’. Basically similar are the BENGALI, ASSAMESE, etc. ‘definitives’ which are suffixed to the noun to make them definite as well as occurring with the numerals as classifiers, e.g. BENGALI *pānc-khana boi* ‘five flat-object book’, *boi-khana* ‘the book’. In the plural definite all classifiers are replaced by the plural *gulo*, e.g. *boi-gulo* ‘the books’. This plural marker cannot occur with numbers.

5.2.2.2. Ordinality

In most languages ordinal numerals are productively derived from cardinal numbers by means of some affix, although it is also true that in many of these languages the first few ordinals are suppletive forms (like English ‘first’ and ‘second’). Some languages (also) employ free elements (e.g. a determiner and/or a relator, which often marks a genitive relation) to express ordinality (Galela, Gude, Nung, Oromo, Samoan; outside the sample also e.g. Chrau, Boumaa Fijian, Modern Greek), and in certain other languages (e.g. Ngiti, Sumerian, West Greenlandic) the expression of an ordinal form requires strategies which are different again. For instance, in Modern Greek ordinals are commonly derived by adding *-to-* or *-(o)sto-* to the cardinal, but in the colloquial variant the ordinals higher than ‘fourth’ often consist of the cardinal numeral preceded by the definite article instead (e.g. *o triánda* ‘the thirtieth’; Joseph and Philippaki-Warburton 1987: 207–8). In Oromo, ordinals higher than ‘twentieth’ must be preceded by the deictic elements *kaa* (masculine) or *taa* (feminine); for ordinals lower than that the various dialects differ as regards the use of these elements (Stroomer 1987: 108). In West Greenlandic ordinals higher than ‘second’ assume the form used for third person plural possession with singular possessum; the head noun is in the relative case (Fortescue 1984: 206–7, 307).

(p.167) In Nivkh the expression of ordinality requires the use of a phrase that includes information about the spatial position of the referent. For example, ‘the first house’ would be expressed as *k'ekr yx p'i dyf*, i.e. ‘the house situated at the end of the upper side of the village’. In the East Sakhalin dialect, however, it seems that several ordinal numerals are formed by suffixing (Gruzdeva 1998: 25).

Ordinal numerals in Ngiti are prenominal relative constructions headed by the verb *isétá* ‘to increase’:

Ngiti (Kutsch Loienga 1994: 358)

(30) ɔyo	ří	kisě	íkyì
two	be	increase	cow
‘the second cow’ [lit. ‘the cow which one adds to be two’]			

Finally there are languages such as Krongo (Reh 1985: 254) and Samoan (Mosel and Hovdhaugen 1992: 319, but see also p. 323) where the difference between cardinality or ordinality is not or not necessarily formally expressed.

In the majority of languages the ordinal numeral occupies the same position as the cardinal numeral. At first sight there are a few languages which do not seem to follow this general tendency, but it must be added that it is doubtful whether we are dealing with grammatical elements that are part of an integral (non-complex) NP. At least in the case of Gude it seems that the ordinal numeral is best analyzed as the head of the whole construction, as is sometimes also the case with cardinal numerals in other languages (see below).

Abkhaz ordinals, which are reported to be of recent origin (Hewitt 1979: 238), must precede the noun, whereas cardinals may precede or follow (see above). Babungo has only one ordinal according to Schaub, but he is reluctant to categorize it as such.²⁵ In Gude the ordinal numeral is categorized as an ‘adjectival noun’ which precedes the noun with which it is linked by the relator *nga* ‘of’ (Hoskison 1983: 51, 56).

Gude (Hoskison 1983: 52)

(31) tufənə	nga	tihinə
fifth	of	horse
'fifth horse'		

(p.168) This category of adjectival nouns also includes nearly thirty elements which would translate as adjectives in English (e.g. ‘white’, ‘damp’, ‘cheap’, ‘expensive’).²⁶

Gude (Hoskison 1983: 56)

(32) nduud'u	nga	la
white	of	cow
'white cow'		

Under an alternative interpretation, which remains undisussed by Hoskison, one could regard the ordinal as the head of a complex NP. In that case the head of the construction would still precede its modifiers (such as cardinal numerals and prepositional phrases).

5.3. Numeral phrases: the lexical expression of the notion Quantity in the noun phrase

The sample contains a number of languages in which cardinal numerals are explicitly categorized as lexical elements. Such languages will be discussed in more detail in this section, which is concerned with the lexical expressions of the notion Quantity in the NP.

5.3.1. Verbal forms

In Koasati (Kimball 1991: 354, 357; see also section 2.2) and Krongo (Reh 1985: 242 f.) numerals are regarded as verbal elements. It may be that in Sarcee, too, numerals are (or perhaps, rather, were) verbal elements: at some point in his discussion of NP complementation Cook (1984: 103) suggests that basic numerals are in fact relativized forms.²⁷

The fact that in some languages cardinals can only be used in a predicative construction does not necessarily mean that they are actually verbs. For instance, in some languages in which the numeral cannot (always) be used attributively, it must occur with a copula. I will not distinguish here between numeral verbs and (p.169) numerals that can only be used predicatively, since in either case the numeral is not a constituent of the simple, integral NP. Thus, in Samoan numerals occur as the head of a special kind of relative clause introduced by the general tense-aspect-mood marker [GENR] *e*; this particular construction is only used as a modifier in a specific NP:²⁸

Samoan (Mosel and Hovdhaugen 1992: 318)

(33) Sa	fau = sia	e	Tagaloaalagi	fale	e	tolu ...
PAST	build = ES	ERG	Tagaloaalagi	house	GENR	three ...
'Tagaloaalagi built three houses ...'						

Alternatively there is also a construction in which the numeral is not the main predicate of an embedded clause but functions ‘as the nucleus of a specific singular noun phrase which is modified by a specific plural possitive noun phrase referring to the counted items’. This construction (which is ignored in Table 5.2) is preferred to the one described above ‘when the counted items are considered as a whole’ (*ibid.* 321):²⁹

Samoan (Mosel and Hovdhaugen 1992: 321)

(34) le	tolu	o	aso
ART	three	POS	day:SP.PL
'three days' (lit. 'the three of the days')			

In non-specific NPs ‘the numeral forms the nucleus of a non-specific singular noun phrase functioning as an apposition’ (*ibid.* 320):

Samoan (Mosel and Hovdhaugen 1992: 320)

(35) ni	popo	se	lima
	ART:NP.PL	coconut	ART:NP.SG five
'five coconuts'			

In Galela numerals must be used predicatively with human nouns, but can be in a direct construction with non-human nouns (van Baarda 1908: 43). The conditions under which the Sumerian numeral is used in a predicative construction are not clear. Thomsen (1984: 83) merely states that ‘sometimes the enclitic copula is added after the numeral’.

5.3.2. Nominal forms

On the whole it is more difficult to establish that a language has numeral nouns. Verbal or predicative numerals are usually easily recognizable, because they (p.170) require a copula, a subject pronoun, or display other properties that clearly prove they are part of a larger clausal construction.³⁰ The reason why numerals are often categorized as nominal lexemes ('nominals', 'nominal words', etc.), however, is simply that they inflect for the same grammatical categories as nouns. This seems to hold for Kayardild (Evans 1995: 85, 233), Nunggubuyu, and West Greenlandic (Fortescue 1984: 204, 247). An additional problem concerning numerals in many Australian languages has to do with their meaning. Hale (1975) has argued that the handful of numerals that are typically found in Australian languages are not numerals in the strict sense but rather indefinite determiners (see also Dixon 1980: 107–8).

Oromo numerals (except 'one' presumably; Stroomer 1987: 106) may occur with or without the suffix *-(aa)ni*. According to Owens (1980: 154) this suffix is related to the plural suffix *-aani*, which I regard as a collective aspect marker. Interestingly, Moreno (1939: 93) calls nouns that occur with this suffix collective numerals ('numerali collettivi'), and translates them into Italian as *i due*, *i tre*, etc. This all suggests that numerals were (and perhaps still are to some degree) nouns marked for collective aspect.

Not much is known about Hurrian, except that the nominal stems for the equivalents of 'two', 'three', 'four', 'seven', 'nine', 'ten', and 'ten thousand' are said to be known to represent Hurrian numerals (Speiser 1941: 82).

In the case of Alambalak, Babungo, and Bambara, all languages in which at least *some* of the cardinal numerals are considered nouns, the evidence is somewhat more substantial, although in the case of Alambalak and Bambara it remains to be proven to what extent the numeral forms in question have retained their lexical character.

Alambalak is said to have only five primitive numbers: *rpa* 'one', *hos* 'two', *tir* 'five' (lit. 'hand/arm') or *wura* 'five' (lit. 'foot/leg'), and *yima* 'twenty' (lit. 'person').³¹ The Alambalak equivalents of 'three' and 'four' are frozen forms which probably derive from a conjoined phrase; they display certain features of nominals in that the form of the person-number-gender (PNG) marker in the NP (-*t* in the example below) is determined by the second root in the cardinal numeral, and not by the head noun:

Alambalak (Bruce 1984: 103)

(36) hos-f-i-rpa	yima-t
two-3DU-CONJ-one	person-3SG.F
‘three men’	

(p.171) The Babungo equivalents of ‘digit(s)’, ‘ten(s)’, ‘hundred(s)’, ‘thousand(s)’, and ‘million(s)’ are categorized as nouns, because these elements all belong to a certain (noun class) gender, just like any other noun (Schaub 1985: 187).³²

Babungo (Schaub 1985: 187)

(37) vé-ŋgá	njò-sé	sé-bò	múu- <u>mbò</u> ø
CL2-antelope	digit-CL	10	CL1O-two ten: PL-two
‘twenty-two antelopes’			

The noun *ɔgá* ‘antelope’ belongs to noun class 1/2, class 2 (CL2) being the plural of class 1 (cf section 3.3.2.2.2 on Bantu noun class systems). The numeral phrase (*njɔ́sè* *sèbò* *múu-mbò*) constitutes a complex noun phrase, involving the nominals *njɔ́* ‘digit’ and *yìghún* ‘ten:SG’. The numeral *-bò* ‘two’ agrees in class with the noun *njɔ́* ‘digit’, which belongs to noun class 9/10 (class 10 is the plural of class 9). This is the class for animals and many other things, such as abstracts. Together *njɔ́sè* and *sèbò* form the numeral phrase expressing ‘two’. The internal structure of the numeral *múu-mbò* ‘twenty’ is less transparent: the plural of *yìghún* ‘ten’ is irregular and its noun class membership unclear (i.e. class 5/?; Schaub 1985: 253–4).³³

One would expect all numbers ending in 1–9 to occur with *njɔ́* (SG) or *njɔ́sè* (PL) ‘digit(s)’, numbers up to 20 to occur with *yìghún* ‘ten:SG’ as well as *njɔ́(sè)* ‘digit(s)’, etc. (thus, in quasi-English, e.g. ‘nine digits’ for 9, ‘nine digits (and) one ten’ for 19, ‘nine digits and two tens’ for 29, etc.). Presumably for reasons of redundancy, the linguistic practice is different: apparently numbers up to 10 do not commonly appear with *njɔ́(sè)* ‘digit(s)’. Similarly, numbers 11–19 usually occur without *yìghún* ‘ten:SG’; in fact, it is added ‘only for clarification and is used mainly by older people’ (Schaub 1985: 238). With numerals higher than 19 Babungo speakers employ the complete construction, i.e. with the numeral nouns *njɔ́(sè)* ‘digit(s)’ and *múu* ‘ten:PL’ (as shown in the example above). Thus, since numbers 1–9 are normally expressed without *njɔ́(sè)* ‘digit(s)’, they look like direct modifiers in a simple NP, whereas in fact they seem to be part of a (headless) phrase modifying the head noun (cf. Rijkhoff 1990a: 23, 30).

Brauner (1974: 41) writes that it is still clear that Bambara numerals were once nouns with a concrete meaning (cf. *waa* ‘1,000’ derives from ‘basket of cola nuts’ → ‘1,000 cola nuts’) ‘so daß ihre Nachstellung zum Substantiv eigentlich eine Genitivverbindung darstellt’ [so that the noun and the numeral are actually in a genitival relationship].

In Hixkaryana numerals are categorized as adverbs and resemble their Australian equivalents in that they are used without great precision (Derbyshire (p.172) 1979: 81, 103–4, 154–5). In Hixkaryana, as in many languages, numerals are now also borrowed

from another language (in this case Portuguese); these numerals are considered nouns in Hixkaryana.

Lango deserves special mention because it remains unclear whether numerals are grammatical or lexical elements.³⁴ What is clear, however, is that (apart from the associative construction) Lango has no grammatical construction which is attributive and does not involve embedding (Noonan 1992: 103, 110; cf. section 4.3.3). Modifiers are all introduced by the ‘attributive particle’ à; in the case of numerals this particle is morphologically fused to the following modifier:

Lango (Noonan 1992: 167)

(38) gúlú àryõ

pot two

‘two pots’

5.4. Conclusion

In this chapter we have seen that adnominal numeral modifiers and number marking are not attested in every language; in fact the sample contains less than ten languages with compulsory number marking. Adnominal modification by a numeral, on the other hand, is possible in most of the languages (although in some of them it seems that the noun modifies the numeral rather than the other way around). In quite a few of them some or all numerals are considered to form a subcategory of the nominal (Babungo) or verbal word class (Samoan). Quantification seems to be the only category that allows for widespread use of non-linguistic modes of expression, since in many languages the cardinality of a referent is still primarily communicated by visual rather than verbal means.

Notes:

(1) Note that Oromo and other languages with set nouns may employ lexicalized forms that contain a nominal aspect marker, which means that semantically the basic lexeme is an individual object noun or a collective noun. As a rule, however, such forms soon become set nouns again (Stroomer 1995: 46; -ca = singulative suffix): ‘Apart from ethnonyms one finds lexicalized forms that contain a singulative suffix, e.g. *billaaca* ‘butterfly’, *d'arti-ca* ‘a liar’ ... In most cases nouns with a singulative suffix lost their singulative meaning and returned to the group of words that are unspecified for number.’ See also the remark below on languages such as Babungo and Kisi, which have overt marking of singular and plural number.

(2) Traditionally Bantu noun classes have been denoted as including number distinctions (section 3.3.2.2.2).

(3) See on reduplication e.g. Moravcsik (1978a); Anderson (1985: 169–70); Mithun (1988: 218 f.); Kiyomi (1995).

(4) See Dressler (1968), Frajzyngier (1985), Durie (1986), Mithun (1988), Newman (1990), and Greenberg (1991b) on number marking with different lexical categories,

Xrakovskij (1989) on iterativity, and Frajzyngier (1997; also Lehmann 1995 and Mithun 1988) on the evolution and grammaticalization of number marking.

(5) But compare Croft (1990a: 99).

(6) See Croft (1990a: 99, 215) for some possible counter-examples, i.e. languages (such as Chumash, Ngandi, Kharia) in which the dual is said to be added to the plural form.

(7) I will ignore Nivkh since numeral classifiers are only used with numerals ‘one’ to ‘five’ (Mattissen and Drossard 1998: 21–2; Gruzdeva 1998: 23).

(8) See also Greenberg (1974: 25) referring to Sanches and Slobin (1973), who found that: ‘Numeral classifier languages generally do not have compulsory expression of nominal plurality, but at most facultative expression.’

(9) On Mandarin *-men*, see also Iljic (1994); Li and Thompson (1989: 40, 83); Chao (1961: 40).

(10) See also Cornyn and Roop (1968: 78): ‘The syllable *-tei* (sometimes *-twei*) in noun expressions denotes the plural. It is much less common than the English plural and is used only when plurality is stressed.’

(11) For more details about the various plural markers in Vietnamese see Luong (1987). Cf. also Vũ Duy-Tù’ (1983: 52) on *các* and *nhữ’ng*, who shows among other things that *nhữ’ng* has the pragmatic function of marking (expanding) focus (cf. Dik 1997: 331–5): ‘Die Verwendung der Substantive im Plural erfordert keine besondere Form: *các* und *nhữ’ng* werden zur Pluralbildung von Substantiven verwendet, wenn klarer hervorgehoben werden soll, daß es sich um eine Mehrzahl von Größen (Menschen, Tieren, Dingen, Angelegenheiten) handelt *các* umfaßt alle Elemente der erwähnten Größe, *nhữ’ng* erfaßt nur bestimmten Elementen davon. ... *nhữ’ng* kann vor einer Zahl verwendet werden, um diese als eine (relativ) große Menge zum Ausdruck zu bringen, z.B. *chiáy có nhữ’ng hai con mèo* “Sie hat sogar zwei Katzen” [To use the noun in the plural does not require a special form: *các* and *nhữ’ng* are used to form plural nouns if one wants to bring out more clearly that a plurality of kinds of entities is involved (people, animals, things, situations) *các* can be employed with all of these entities, *nhữ’ng* only with some of them. ... *nhữ’ng* can be used before a numeral ... to express a (relatively) large number, for instance, *chiáy có nhữ’ng hai con mèo* [she (lit. older sister) DEM have PL two CL:animal cat—JR] ‘she even has TWO cats’].

(12) See e.g. Menninger (1969); Bartsch (1973); Stampe (1976); Greenberg (1978b); Hurford (1987); Seiler (1990); Gvozdanović (1992, 1999); Comrie (1997); Heine (1997: 18–34).

(13) For the numbers 1 to 4, however, Kobon also has alternative forms.

(14) According to Greenberg (1966a: 86) doubling of the numeral also occurs in Guarani; on doubling of numerals see also Greenberg (1978b: 284) and Greenberg (1989).

(15) Grammatical relations between ‘numerals’ and the head noun are not always clear in the Australian languages (cf. Merlan 1982: 92–3 on Mangarayi; Donaldson 1980: 225–6 on Ngiyambaa; and Heath 1984: 497 f. on Nunggubuyu); often the numeral (if it is that; see below) is thought to have an appositional relation with the noun.

(16) If Greenberg (1989) is correct, the position of the lowest numerals reflects the older pattern.

(17) See e.g. Hagman (1974: 90 f.) on appositives in Nama Hottentot.

(18) In Nung the classifier phrase precedes the noun, but one variant of ‘one’ follows (see above).

(19) See also Lee (1989: 50–2, 59) on Korean. Korean has two sets of numerals: ‘Korean numerals’ and ‘Sino-Korean numerals’. The latter set is used for numbers from 100 onward. Korean numerals collocate normally with the pure Korean nouns, whereas numerals from the other set occur in noun phrases with nouns of Chinese origin. It seems that at least some of the Korean derived numerals may occur before the head noun without being accompanied by a classifier, as in *se ceg* (*se* from Korean *ses* ‘three’) [three book] ‘three books’.

(20) On numerals in Nivkh, see also Gruzdeva (1998: 24) who writes: ‘The syntactic peculiarity of the numerals up to “five” is that in counting phrases they are in postposition to the counted nouns and are declined according to the rules of noun declension (with the head noun in the nominative) ..., whereas all other numerals are always placed in preposition ..., and are declined only in the isolated position.’

(21) In many languages numeral classifiers have a variety of other functions (section 3.3.1.2). It remains to be seen to what extent this correlates with their position relative to the noun (Adams et al. 1975: 2; see also Adams 1986: 224–45; Greenberg 1972: 6; Goral 1978; Barron and Serzisko 1982: 97–9): ‘Besides their function in numeral noun phrases classifiers in various languages function as nominal substitutes, nominalizers of words in other form classes, markers of definiteness, relativizers, markers of possession, and as vocatives; serve to disambiguate sentences, establish coherence in discourse and regularly mark registers and styles within a language.’

(22) This may already happen when the classifier phrase follows the noun (Jones 1970).

(23) Consider also this remark by Greenberg (1975: 30): ‘Many classifier languages have constructions in which the classifier appears without a quantifier. The meaning here is invariably singular, and is in some languages specified as definite, in others indefinite and still others neither. In some grammars it is explicitly noted that this occurs through deletion of “one” (the most unmarked number).’

(24) Cf. also Lehmann (1982b: 255).

(25) Cf. Schaub (1985: 240–1): ‘It is questionable whether this word [the equivalent of “first”] should be classified as an ordinal or as a modifying adverb or emphasis word, because it is invariable. (It does not agree in noun class with the head-noun.) In addition, it does not occur in the usual position of numerals but precedes the noun. (Note that ordinarily only emphatic words precede the noun; ...) With all other numerals there are no ordinal numbers. Cardinal numbers have to be used, and the differentiation may be made by the verbs ... “fill” or “follow”: ... Often, however, people are inexact and do not differentiate. For example they may just say “I bought two

books” when they bought the second book. In this case it may be known only from the context whether the speaker means “the second book” or “two books”.

(26) See also Rischel (1995) on a similar construction involving adjectives in Minor Mlabri (for a general discussion of adjectives with possessor nouns or ‘dependency reversal’ see Ross 1996 and Malchukov 2000). The problem of headedness also occurs in Koasati (Kimball 1991: 484) and in the Australian languages, where a (specific) noun is often preceded by generic noun, as in: *kiyarrng-ka yarbud-a ngarnal* [two-NOM meat-NOM white_cockatoo:NOM] ‘two white cockatoos’ (see 3.3.1.1). Here I follow Evans (1995: 235; see also Wilkins 2000: 151) and regard such constructions as appositional nouns which together form a composite head (see also Corbett and Fraser 1993 on the notion of head in grammatical theory).

(27) See also e.g. Holmer (1947: 104) on Cuna, where ordinal numerals may be expressed by suffixing the participial forming *-t(i)* to the cardinal form.

(28) In his grammar of Boumaa Fijian (a language that is closely related to Samoan), Dixon (1988: 144) analyzes *e* as a third person subject pronoun, as in: *e tolu a gone* [3SG.S three ART child(ren)] ‘three children’.

(29) I am ignoring constructions in which the numeral is part of a compound (Mosel and Hovdhaugen 1992: 322–4).

(30) In theory a numeral noun could also serve as the main predicate of a clausal construction, but so far I have found no such cases.

(31) Note that in Alamblak the equivalent of ‘two’ may itself occur with the dual marker (see also Merlan 1982: 86 on Mangarayi and Merlan 1983: 53 on Ngalakan). I assume that numerals which may have their own ‘number’ marker are in fact nouns (Corbett 1978: 358; Hurford 1987: chapter 5). Number (or aspect?) marking on numerals is (also) attested in other languages of the Cushitic family (Hetzron 1967). For instance, Sidamo decimals from ‘fifty’ to ‘ninety’ are the plurals of digits from ‘five’ to ‘nine’ (Greenberg 1978b: 279; he also has examples from Semitic languages).

(32) Greenberg (1989: 111) writes: ‘The lower a numeral is, the more it is treated as an adjective and the higher it is ... the more it is treated as a noun, morphologically and syntactically’ (see also Corbett 1978).

(33) There is an increasing tendency among Babungo speakers, especially the younger people, to use English numerals (Schaub 1985: 239).

(34) Lango is a Nilotic language of the East Sudanic branch of Nilo-Saharan. In the grammar of Nandi, another member of the Nilotic family, it is simply stated that numerals are morphologically nouns (Creider and Creider 1989: 64; see also Chapter 9).

6 Localizing Modifiers in the Noun Phrase

6.1. Introduction

Besides nouns, we have dealt with two kinds of NP constituents so far, namely

1. Qualifying operators and certain qualifying satellites (nominal aspect markers and adjectives; Chapter 4), so labeled because they have to do with qualitative properties of the referent of the NP;
2. Quantifying operators and quantifying satellites (number markers, numerals, and numeral phrases; Chapter 5), so called because they pertain to quantitative properties of the referent of the NP.

This chapter is concerned with the third major modifier category in the NP: *localizing operators* and *localizing satellites*, which relate to locative properties of the referent of the matrix NP. Examples of such constituents are:

<u>Localizing operators in NP</u>	<u>Localizing Satellites in NP¹</u>
demonstrative pronoun	(localizing/identifying) relative clause
(in)definite article	possessive modifier locative modifier

At first sight it may not be immediately clear what some of these noun modifier categories have to do with the notion ‘location’, but a closer examination will reveal that each of these constituents typically indicates that the referent of an NP has a place in the world of discourse (or that this place can be inferred; for example, the indefinite article usually indicates for the hearer that the referent of the NP is *introduced* into conversational space).²

This localizing function is rather obvious in the case of demonstrative pronouns (*this book*) and locative modifiers (*the book on the table*), which both (p.174) rather straightforwardly specify the spatial location of the referent in question.³ However, spatial references often serve as the basis for metaphorical extension into other domains (Lakoff and Johnson 1980), so that constituents indicating distance/location in space may eventually be employed to indicate distance/location in time (temporal space) and ultimately also in pragmatic, discourse, or psychological space.⁴ Thus deictic elements such as demonstrative pronouns may come to be employed as anaphoric elements and in many languages there is a synchronic or diachronic relationship between demonstratives on the one hand and definite articles and third person pronouns on the other.⁵ But whereas (deictic) demonstratives are primarily used in connection with the location of a referent (or rather its real-world counterpart) in the physical world, definite articles are more commonly used in connection with referents that have already been mentioned in the discourse or that otherwise are assumed to be identifiable (by inference, general knowledge, or situational conditions; see below on identification). In other words, referents of definite NPs are supposed to be *locatable* in the ongoing discourse or in the knowledge base of the speaker (I am ignoring terms with generic or specific reference, which require special treatment; cf. sections 1.5.4 and 7.7).

It is perhaps less obvious how relative clauses and possessive modifiers can serve as localizing constituents in the NP. To start with the latter, it has been shown in various studies that cross-linguistically there are some remarkable similarities between locative and possessive constructions (as well as existential constructions, which are not so relevant here; see also Chapter 2). Examples of such constructions are (Lyons 1967: 390; Clark 1978: 86; Christie 1970):

- | | |
|----------------------------------|---------------|
| (1) There is a book on the table | existential |
| (2) The book is on the table | locative |
| (3) John has a book | possessive I |
| (4) The book is John's | possessive II |

Such parallels were pointed out by Lyons (1967), who was the first to suggest that these constructions are related, both synchronically and diachronically. The nature of this relation was investigated in more detail by Clark (1970, 1978), who demonstrated on the basis of a sample of sixty-five languages that these constructions are systematically connected in terms of word order and patterns of verb(p.175) use. The rationale behind these similarities is that the constructions under consideration all contain what she called *locationals*. In the first two examples this *locational* is the locative modifier ‘on the table’; in the last two it is the possessor ‘John(‘s)’. Clark interprets possessors as locations, saying that, cognitively, possessed items can be argued to be located ‘at’ the possessor (Clark 1970: 3):⁶

psychologically it would appear quite plausible to argue that if an object is *in some place*, and the ‘place’ is actually an animate being, then the object is *possessed by* the ‘place’. In other words, it is the [+Animate] feature added to the locative phrase that transforms it into a Possessor-nominal.

Although Lyons and Clark were only concerned with sentential constructions, this localist interpretation of possession also holds for adnominal possessive modifiers, of course, whether this involves a possessor noun or a possessor pronoun. An example of a locative element that has developed into a marker of possessorship can be found in this example from Ewe (Niger-Kordofanian).

Ewe (Claudi and Heine 1986: 316)

(5) fofó	nye	φé	xo
father	my	place	house

‘my father’s house’ (‘the house at my father’s place’)

According to Claudi and Heine (ibid.), ‘the relational noun φé “place” was used as a vehicle to denote possession and developed into a general marker of nominal possession’.

Turning now to relative clauses, these often serve to give the referent of the relativized NP a place in the temporal dimension.⁷ That is, relative clauses designate events, which are located in time, and since the referent of a relativized NP is necessarily a participant of the event denoted by a relative clause, this participant is located in time (and space) as well. Thus ‘the book’ in ‘the book I lost in the train yesterday’ is located in the past by virtue of it being a participant in the event denoted by the relative clause in the past

tense, which also contains the adverbial ‘yesterday’, giving an even more precise temporal indication.

Let me finally point out that localizing satellites play an important role with respect to the identifiability of the referent of the matrix NP.⁸ This is because localizing satellites are the constructions that contain the NP through which the referent of the matrix NP can be *anchored* in conversational space (see below). The name for this discourse strategy is due to Prince (1981: 236), who describes (p.176) it as follows: ‘A discourse entity is Anchored if the NP representing it is LINKED, by means of another NP, or “Anchor”, properly contained in it, to some other discourse entity.’⁹ To put it differently, localizing satellites make it possible to introduce a new referent through a definite NP, allowing the speaker to present a new referent as an identifiable entity. Consider these examples (* = pragmatically marked):

- (6) *Tomorrow I'll buy *the book*.
- (7) *I went to see *the apartment last night*.
- (8) *I went to *the restaurant* to have lunch.

Each of these sentences contains definite NPs (in italic print), suggesting that their referents are assumed to be identifiable for the hearer, whereas in fact they are not (unless under special circumstances). In each case the hearer has the inclination to ask ‘What book?’, ‘What apartment?’, or ‘What restaurant?’ Now compare these sentences:

- (9) Tomorrow I'll buy *the book you want so much*.
- (10) We went to see *John's apartment last night*.
- (11) We went to *the restaurant in the Van Gogh Museum* for lunch.

This time each NP is provided with an (underscored) embedded localizing satellite (i.e. the *anchor*), which enables the hearer to properly identify the referent of the matrix NP. Now the hearer can identify ‘the book’ as ‘the book he or she wanted so much to have’. In a similar vein ‘the apartment’ can now be identified as ‘the one occupied by John’ (with whom the hearer is presumed familiar), mentioned in the possessive modifier; and ‘the restaurant’ can now be identified as ‘the one in the Van Gogh Museum’ (also assumed known), referred to in the embedded locative NP.¹⁰

Anchoring is in fact part of a more general identification procedure, which says that a referent can only be identified by linking it with its *Identifying Entity* (IE). Exactly what will serve as an IE depends on various factors, but an IE must itself always be something that can be identified. It may be the referent of an NP contained in a localizing satellite, as in the examples above (‘you’, ‘John’, ‘the Van Gogh Museum’, respectively). But in the case of demonstratives (in their deictic function) the IE is the object of the same name in the physical world, and in the case of a definite article (used anaphorically) the IE is the referent that was construed (p.177) earlier and is stored in the referent file in the knowledge base of the speech participants (cf. Heim 1982; for a more detailed discussion and illustration of this general identification procedure I refer the reader to Rijkhoff 1989).

The identifying function of localizing satellites also explains why they tend to occur in definite NPs, since it is (the reference to) the IE in the localizing satellite that often

licenses the definiteness of the matrix NP. In fact there are languages in which relative clauses are almost exclusively attested in definite NPs. Consider in this context, for example, these remarks from Lehmann's (1984) monograph on relative clauses (see also Moravcsik 1969: 167, 170; Bach 1974: 192, 272; Givón 1990: 645 ff.).¹¹

Mit einem Relativsatz kann man leicht einen bestimmten Gegenstand durch Spezifikation der Situation, an der er teilhat, identifizieren. So erklärt es sich, daß die typische Relativkonstruktion von einem Definitum begleitet ist, wiewohl das natürlich prinzipiell nicht notwendig ist.

[With a relative clause one can easily identify a certain object by specifying the situation in which it is involved. This explains that the typical relative construction co-occurs with a determiner, although in principle this is not necessary, of course.] (Lehmann 1984: 402)

Das Adjektiv dient mehr der Begriffsbildung, der Relativsatz mehr der Gegenstandsidentifikation.

[The adjective primarily adds to the meaning, the relative clause typically serves to identify an object.] (ibid. 405)

It seems that the possessive modifier also lends itself quite easily for identifying purposes, since the possessor NP is often found in multiple embeddings. Here is an example from Babungo (used earlier in section 1.5.4.3) which involves a fourstep identification procedure. In this example *tó* ('head'/‘hair’) is identified through ‘Lambi’, then ‘Lambi’s wife’, then ‘Lambi’s wife’s brother’ and finally ‘the child of Lambi’s wife’s brother’:¹²

Babungo (Schaub 1985: 76)

(12) (ŋwé bá)	tó	wée	wéenshú	zú	Làmbí
(she weave.PRES)	head	child	brother	wife	Lambi
‘(she is plaiting) the hair of the child of Lambi’s wife’s brother’					

(p.178) As to NPs functioning as localizing satellites with such semantic functions as Beneficiary ('the flowers *for Mary*') or Agent ('the article *by Smith*'), matters are perhaps somewhat less clear. At this point it may be relevant to refer to the localist hypothesis, according to which language is built on a spatial metaphor in that spatial expressions are thought to be more basic than various kinds of non-spatial expressions (Lakoff and Johnson 1980: 14 f., 56 f.; but cf. Heine et al. 1991: 114–18).

Spatial expressions are linguistically more basic, according to the localists, in that they serve as structural templates, as it were, for other expressions; and the reason why this should be so, it is plausibly suggested by psychologists, is that spatial organization is of central importance in human cognition. (Lyons 1977: 718)

Thus, ultimately, non-spatial semantic functions like Beneficiary and Agent are conceptualized in terms of spatial notions, which would explain why (originally) locative elements were used to express non-locative, non-spatial semantic functions (Clark 1978: 114). In fact, Hjelmslev (1935) and Anderson (1971) have argued that nearly all adpositions are locative in origin. This theory is substantiated in numerous

studies which demonstrate how non-spatial meanings have developed out of spatial elements.¹³ Next we will be looking at *localizing operators* (6.2) and *localizing satellites* (6.3).

6.2. Demonstratives and articles: the grammatical expression of the notion Location in the noun phrase

6.2.1. Demonstratives

Demonstrative pronouns primarily serve to indicate the relative distance between the real-world counterpart of a referent and a certain reference point, which usually coincides with the speaker's position. Demonstratives may also encode information pertaining to such diverse phenomena as visibility, shape, height (relative to the speaker), and geographical or environmental features.¹⁴ It has already been mentioned that spatial references often serve as the basis for metaphorical extension into other domains, so that eventually demonstrative pronouns may become, for example, definite articles, personal pronouns, number markers, tense markers, or complementizers.¹⁵

In the present context I am mainly interested in those morpho-syntactic properties of demonstratives which are relevant in connection with the ordering (p.179) principles to be discussed in Chapters 8–10. Table 6.1 presents the position(s) of the attributive demonstrative pronoun (or demonstrative adjective) relative to the head noun. Notice that in some languages the demonstrative may or must occur with an article (ART) or a classifier (CLF). This as well as certain other peculiarities is discussed below.

Table 6.1. The expression of attributive demonstratives

Abkhaz	dem ART N
Alamblak	dem N dem
Babungo	dem N dem
Bambara	dem N dem <u>or</u> : dem N-in
Basque	N dem
Berbice Dutch Creole	ART N dem
Bukiyip	dem N dem
Burmese	dem N
Burushaski	dem N
Cayuga	(see Table 6.2 note 2)
Chinese, Mandarin	dem N <u>or</u> : dem CLF N
Chukchi	dem(-) N
Dutch	dem N
Galela	ART N dem
Georgian	dem N
Guaraní	dem ART N
Gude	N -dem

Hittite	dem N (ex.)
Hixkaryana (see below)	
Hmong Njua	CLF N dem
Hungarian	dem ART N
Hurrian	N dem
Ika	dem N
Kayardild ¹	dem N
Ket	dem N
Kisi	N dem
Koasati	dem N
Korean	dem N
Krongo	N dem
Lango	N (-) dem
Nahali	dem N
Nama	dem N
Nasioi	dem N dem
Ngalakan	dem N
Ngiti	dem N dem
Nivkh	dem(-) N
Nung	N dem
Nunggubuyu	dem N dem
Oromo	N dem
Pipil	dem N
Quechua, Imbabura	dem N
Samoan	ART-dem N <u>also e.g.</u> ART N dem (see below)
Sarcee	dem N
Sumerian	N (-) dem
Tamil	dem N
Tsou	N dem
Turkish	dem N
Vietnamese	N dem
Wambon	dem N
West Greenlandic	dem N dem

(1) Kayardild ‘the demonstrative *dathina* “that” doubles as a discourse determiner’ (Evans 1995: 86) and ‘all modifiers precede the head, except that one modifier may be postposed’ (*ibid.* 235).

(p.180) Not every language is reported to have attributive demonstrative pronouns. This is most clearly stated in the case of Hixkaryana: ‘Demonstrative pronouns are not used adjectivally’ (Derbyshire 1979: 131). They do occur as distinct terms by themselves, as in this equative construction.

Hixkaryana (Derbyshire 1979: 132)

- | | |
|-----------------------|--------------------------|
| (13) a. romuru mosoni | b. Waraka kanawari moro |
| my-son this-one | Waraka canoe-of that-one |
| ‘This is my son’ | ‘That is Waraka’s canoe’ |

In certain other languages, like those spoken on the Australian continent, it is often suggested or stated that the demonstrative is not used attributively either. Instead it seems to have an appositional relation with another term denoting the same referent (see Donaldson 1980: 138, 229–30, 232 on Ngiyambaa; Heath 1984: 501–2 on Nunggubuyu).¹⁶ In West Greenlandic (Fortescue 1984: 110) and perhaps in Nasioi (see Chapter 9), demonstratives can also be regarded as appositional elements.

6.2.1.1. Bound demonstrative forms

In the following languages the demonstrative may or must be expressed as a bound form (clitic or affix): Chukchi, Gude (Hoskison 1983: 45), Lango, Nivkh **(p.181)** (Mattissen and Drossard 1998: 19–20) and Sumerian.¹⁷ Chukchi and Lango have both a free and a bound variant. Apparently speakers of Lango normally use the bound form, but the forms that are used as independent demonstrative pronouns ‘may also substitute for the demonstrative affixes’ (Noonan 1992: 165). In Chukchi the free form is only used in the absolute (see also Bogoras 1922: 724 on the role of emphasis).

Lango (Noonan 1992: 165)

- | | |
|------------------------|-----------|
| (14) òkélò òkwàñò | búkki |
| Okelo 3SG:read:PERF | book-this |
| ‘Okelo read this book’ | |

Lango (Noonan 1992: 166)

- | | | |
|------------------------|------|------|
| (15) òkélò òkwàñò | búk | mân |
| Okelo 3SG:read:PERF | book | this |
| ‘Okelo read this book’ | | |

In Sumerian the situation is not altogether clear (Thomsen 1984: 80):

In Sumerian there is a number of demonstrative pronouns and suffixes, but they are not used very often, and the semantic distinctions between the various demonstratives are therefore not clear, *ne.en* or *ne.e* ‘this’, is perhaps opposed to *ri* ‘that, yonder’, which, however, is limited to fixed literary expressions ... The occurrences of *sé* and *-e* are doubtful, and *-bi* which is often used in the sense ‘this’, is simply the possessive suffix of inanimate and collective.

6.2.1.2. Doubling of demonstratives

In a number of languages the demonstrative may appear on either side of the head noun. In Alambalak, for instance, demonstratives normally precede the noun but the reverse order was only attested in elicited speech (Bruce 1984: 100). In Babungo demonstratives normally follow the noun, but the emphatic form, which can also constitute a term by itself, may precede (Schaub 1985: 73, 204 f.). Conversely, demonstratives in Bambara normally occur before the noun, but in rare cases they are also placed in postnominal position:

Bambara (Brauner 1974: 48)

- | | |
|------------------|--------------|
| (16) a. nin muso | b. dugu nin |
| this woman | place this |
| 'this woman' | 'this place' |

(p.182) For emphasis, the noun (or adjective) may carry the *-in* suffix

Bambara (Brauner 1974: 48)

- | | |
|-------------------|------------------------|
| (17) a. nin cè-in | b. nin kulu ba-in |
| this man-IN | this mountain big-IN |
| 'this man' | 'this big mountain' |

A Bukiyp adnominal demonstrative also normally precedes the noun, but if it is the only optional noun modifier, it may also appear after the noun (Conrad 1991: 57). Demonstratives in Ngiti normally appear before the noun, too, except in a question with the question word in clause-initial position (SC = subject concord), compare:

Ngiti (Kutsch Lojenga 1994: 374)

- | | |
|--|-----------|
| (18) wòkć ibile dza | |
| those three house | |
| 'those three houses' | |
| | |
| (19) ádhà nyòdzì | mbérù yà? |
| where SC.2SG-buy:PERF.PRES clothes | these |
| 'Where have you bought these clothes?' | |

6.2.1.3. Demonstratives co-occurring with articles

In the following languages the demonstrative may or must co-occur with what is usually called an article: Abkhaz, Berbice Dutch Creole, Galela, Guaraní, Hungarian, and Samoan.¹⁸

At least in the case of Abkhaz and Galela this phenomenon may be understood in the light of a theory of diachronic change proposed by Greenberg (1978a: 65, 1981), according to which definite articles (in Greenberg's terminology: Stage I articles) may

develop into noun or gender markers (or Stage III articles). This theory was discussed in section 3.3.2.4.1.

The so-called articles in the aforementioned languages display properties which, according to Greenberg's theory, are typical for articles that are in an intermediate phase in this diachronic process (i.e. in Stage II). The most characteristic property of the articles at this point in their development is that they occur both in definite and in indefinite NPs, or, more generally, in NPs with specific reference.¹⁹

(p.183) The Abkhaz article is prefixed to nouns (or preposed adjectives of nationality, if any). Not only does it occur in definite and indefinite NPs, it also appears on nouns in their citation forms (another characteristic of Stage II articles).

Abkhaz (Hewitt 1979: 57, 225)

(20)	wəy	à-jxab
	that_one	ART-girl
'that girl'		

(21)	à-kərt + wa	jxab
	ART-Georgian	girl
'the/a Georgian girl'		

In Samoan the NP usually begins with what is called an article (only plural NPs with non-specific reference lack one), but which could also be characterized as a noun phrase marker (Crowley 1985). The article does not mark (in)definiteness, but rather (non)specificness and also occurs in generic statements (Mosel and Hovdhaugen 1992: 259):

Samoan (Mosel and Hovdhaugen 1992: 259)

(22)	ε	ai-na	le	gata
	GENR	eat-ERG	ART	snake
'Snakes are edible'				

Demonstratives, which may occur on either side of the head noun, are bound forms when they follow the singular article:

Samoan (Mosel and Hovdhaugen 1992: 291)

(23)	lea	tama
	that	boy
'that boy'		

(24)	le	'upu	lea
	ART	word	that

'that word'

-
- (25) si-nā pua 'a

ART-that Pig

'that poor pig' (note that si has an emotional quality)

-
- (26) si teine si-a

ART girl ART-that

'that (dear little) girl'

According to Mosel and Hovdhaugen (1992: 292), postnominal article-demonstrative combinations that involve the specific-singular article *le* (namely *le-nei*, *le-nā*, *le-lā*, *le-a*) must be regarded as appositions.

(p.184) Samoan (Mosel and Hovdhaugen 1992: 292)

-
- (27) 'o le tama le-a

PRES ART boy ART-that

'the boy, that one'

Galela has a 'general article', which occurs in both definite and indefinite NPs and which also seems to function as a noun or noun phrase marker (cf. also Holton 1999: 346 and Visser and Voorhoeve 1987: 36–8 on noun markers in two related languages, Tobelo and Sahu). For example (van Baarda 1908: 33, 51):

Galela (van Baarda 1908: 51)

-
- (28) o tahu manèna

ART house this

'this house'

The same element appears to function as a nominalizer; compare:

Galela (van Baarda 1908: 18)

-
- (29) hau

'to fish'

-
- (30) o hau

ART fish

'the (manner of) fishing'

Examples indicate that in Guaraní the article sometimes co-occurs with the demonstrative, but it is not explained under what circumstances this happens (Gregores and Suárez 1967: 147, 150):

Guaraní (Gregores and Suárez 1967: 147, 150)

- (31) upé (la) ?óga
that (the) house ...
'those houses ...'

6.2.1.4. Demonstratives co-occurring with classifiers

In three languages in the sample the demonstrative occurs with a numeral classifier: Hmong Njua, Mandarin Chinese, and Vietnamese.

Mandarin Chinese (Li and Thompson 1989: 105)

- (32) nèi-tiáo niú
that-CLF cow
'that cow'

Hmong Njua (Harriehausen 1990: 140)

- (33) phau ntawv nuav
CLF book DEM
'this book'

(p.185) Vietnamese (Nguyễn Đình-Hoá 1987: 785)

- (34) con bò ấy
CLF cow that
'that cow'

It is perhaps not accidental that in all three languages the classifier has assumed other or additional functions (it was mentioned in section 5.2.2.1.3 that the classifier often takes on other functions when it appears in prenominal position). In Chinese the classifier only occurs in terms with specific reference (Li and Thompson 1989: 130; Sun 1988), in Hmong Njua it is used to express nominal aspect (section 4.2.1.2), (in)definiteness, size, and appearance (Miao Language Team 1972: 14), and in Vietnamese it marks definiteness (Löbel 1999: 310).²⁰

6.2.2. Definiteness and indefiniteness markers

Definite articles can also be regarded as localizing elements, albeit that with these elements (as in the case of personal pronouns and temporal deictics) we rather speak of 'weak deixis' (Anderson and Keenan 1985: 261–2). This section focuses on the form and position of elements in the NP that mark definiteness or indefiniteness, i.e. elements that indicate whether or not the referent of the phrase in question is regarded as being

identifiable for the hearer in the linguistic or non-linguistic context.²¹ Indefiniteness markers are included as well because they are used to introduce a referent into conversational space (Chapter 7).

It is important to emphasize that I do not take into account modes of expression outside the NP, as when, for instance, only definite NPs are cross-referenced in the predicate. Neither do I consider special syntactic properties of definite or indefinite NPs. For instance, in many languages only definite terms will appear in a sentence initially (see e.g. Fortescue 1984: 182 on West Greenlandic).²²

Definiteness and indefiniteness are expressed in a variety of ways. Many languages have no special (in) definiteness markers; if necessary a (distal) demonstrative pronoun or the numeral ‘one’ may be used to express the notions of definiteness and indefiniteness respectively, as in e.g. Mandarin Chinese (Li and Thompson 1989: 85–6) and Tamil (Asher 1982:61). Other languages have separate (p.186) forms that are clearly diachronically related to the distal demonstrative (Laury 1997) or the numeral ‘one’ (Givón 1981a), and yet another group of languages employs expressions where this relationship is not so obvious or lacking altogether.²³ In the latter case the definite article appears to have developed from a variety of other sources, such as numeral classifiers (see below), positional verbs (see e.g. Watkins 1976b: 4; Rankin 1977; Moser 1977: 20–1; Barron and Serzisko 1982), or participializing or nominalizing elements (see below). The sample even contains a language (Nasioi) in which the definiteness marker appears to have developed from the numeral ‘one’.

There are also languages in which (in)definiteness is expressed by elements that may encode other notions as well, as when the (in)definiteness marker is contained in a portmanteau element which simultaneously expresses gender/class, number and/or case.

Table 6.2 indicates how (in)definiteness is expressed in the NP. If anything, this table clearly shows that for many languages it is not compulsory to indicate that an NP is definite or indefinite, at least not in the NP itself (a minus (-), (p.187)

Table 6.2. The expression of (in)definiteness

Language	Def	Indef	
Abkhaz	(section 6.2.1.3)	N-ind	one
Alamblak	(dem) N	N-ind?	indef.gender
Babungo	N dem	Ø	
Bambara	Ø	Ø	
Basque	N-def	N ind	one
Berbice Dutch Creole	def N	ind N	one
Bukiyip	–	–	
Burmese	–	–	
Burushaski	dem N ¹	N-ind?	one
Cayuga ²			
Chinese, Mandarin	(dem-clf) N	(one-clf) N	

Chukchi	N-def?	—	
Dutch	def N	ind N	one
Galela	(section 6.2.1.3)	Nind	one
Georgian	Ø	Ø	
Guaraní (but see above)	def N	ind N	one
Gude	N-def	—	
Hittite	—	—	
Hixkaryana	Ø	Ø	
Hmong Njua	elf N	elf N?	
Hungarian (but see above)	def N	ind N	one
Hurrian	—	—	
Ika	(dem) N	?ind N	one
Kayardild (see above)	(dem) N (dem)	—	
Ket	?	?	
Koasati	N-def	—	
Korean	(dem) N	—	
Krongo	—	—	
Lango	Ø	N-sx?	
Nahali	—	—	
Nama Hottentot	(dem) N	N-ind?	indef.gender
Nasioi	N-def	Ø	
Ngalakan	(dem) N	Ø	
Ngiti	—	—	
Nivkh	(dem) Ø	Ø	one
Nung	—	—	
Nunggubuyu	(see below)	—	
Oromo	N (dem)	N ind	one
Pipil	def N	ind N	
Quechua, Imbabura	Ø	Ø	
Samoan (see above)			
Sarcee	—	—	
Sumerian	—	—	
Tamil	(dem) N	ind N ind	one
Tsou	Ø	Ø	
Turkish	Ø	Ø	one
Vietnamese	Ø	Ø	
Wambon	—	—	

(1) See Berger (1998: 87) on demonstratives serving as definite articles in Burushaski; cf. also Berger (1998: 43–4) and Lorimer (1935–8: i. 17) on the function of the so-called determinative or definitive suffix.

(2) In Cayuga (Mithun 1992: 57) '[t]he relative order of determiners and nominals is generally invariable. Definiteness is not obligatorily marked in Cayuga, and no indefinite article is used. ... an optional particle *ne'* may precede definite nominals, including proper and possessed nouns. When it appears, it precedes the nominal it modifies. This position is, of course, functional. In Cayuga, morphological verbs, like any clauses, can function as syntactic nominals. They need carry no overt markers of nominalization. The result is that normal discourse can consist largely of verbs. The particle *ne'* is most often used to indicate that what follows is functioning syntactically as a nominal.'

Cayuga (Mithun 1992: 57)

(2)	akaqnihantó:k	ne'	nóne:'	né: kyé	ne'	kowiyáqtatre'
	they.noticed	the	you.know	this	the	she.is.getting.a.baby
'... they noticed that she was expecting'						

Notes: Ø = language is stated not to have (the equivalent of) (in)definite articles; – = no mentioning of (in)definiteness markers; question marks are explained below in section 6.2.2.1.2.

(p.188) which means that no information on the expression of (in)definiteness was found, is probably often simply an indication that these notions receive no distinct expression in the language in question).

Notice that I do not regard every element that is called ‘article’ in the literature as an instantiation of an (in)definiteness marker. Let me in this context refer to e.g. sections 3.3.2.4.1 and 6.2.1.3, where it is argued that certain so-called articles are in fact (in the process of becoming) noun or gender markers, i.e. Greenberg’s Stage II and Stage III articles.

6.2.2.1. The expression of definiteness

6.2.2.1.1. Demonstratives

Often definiteness markers are treated on a par with attributive demonstrative pronouns, as when both categories are subsumed under the label ‘determiner’. However, morpho-syntactic (and other) properties of demonstratives and articles need not necessarily coincide (Dryer 1989a: 90; Van Valin and LaPolla 1997: 62–3). In Berbice Dutch Creole, for instance, *di* is used as both a definiteness marker and a proximate demonstrative, but the former precedes and the latter follows the noun:²⁴

Berbice Dutch Creole (Kouwenberg 1991: 111)

(35)	di	wari	di
	the	house	this

'this house'

One should also distinguish between definite and indefinite articles, for that matter, since they may also take up different positions in the NP (Moravcsik 1969: 86; Krámský 1972).

Nevertheless, it is true that in some languages a demonstrative form tends to be used as a definiteness marker, as in the Berbice Dutch example above. In such cases the demonstrative has usually lost at least some of its purely deictic function.²⁵ For example, the Oromo demonstrative *kani* 'this, these' (Boraana dialect) also occurs with an internal long vowel (*kaani*), in which case it means 'this, that, these, those, the other one, the one mentioned before':

Boraana Oromo (Stroomer 1995: 63)

(36)	worrii	sangaa	kaani	k'ale
	people:S	ox	the_one_mentioned_before:M	slaughter:3SG.M.PAST
'They slaughtered the ox mentioned before'				

(p.189) In Nunggubuyu, too, definiteness can be expressed by a demonstrative pronoun (Heath 1984: 270). However, there is evidence to suggest that demonstrative pronouns are often in an appositional relation with the head noun (ibid. 505; see also below on appositional definiteness markers in Ngiyambaa).

6.2.2.1.2. Distinct definiteness markers

In quite a few languages definiteness appears to be expressed by some special element. In Basque, possibly Chukchi (see note 26), Gude, Koasati, and Nasioi this element is a bound form, whereas in Dutch, Guarani, Hungarian, and Pipil it is a free form.²⁶ The question mark after some of the languages in Table 6.2 already indicates that it is sometimes difficult to decide whether or not one may actually speak of a definiteness marker, as in the case of some element whose function is not quite clear and which only tends to occur in an NP referring to an identifiable entity.²⁷

Elements that appeared to function as topic markers were excluded as well (cf. Noonan 1992: 161 on Lango *-mère*). Although there is a certain correlation between topicality and definiteness in that the topic of the sentence is commonly referred to by means of a definite NP, the reverse is not necessarily true (Dik 1997: chapter 13). Here are some examples of bound definiteness markers.

Basque (Saltarelli 1988: 81)

(37)	gona	gorri	estu-ak
	skirt	red	tight-PL.ABS
'the tight red skirts'			

In Basque the definiteness marker is more or less amalgamated with the 'number'/case suffix: *-a* for definite singular NPs, *-ak* for definite plural absolute NPs, and *-e* for

definite plural ergative NPs; indefinite NPs are unmarked (Saltarelli 1988: 198; see section 4.2.1.2 on ‘number’ in Basque).

Gude uses a suffix to mark definiteness:

Gude (Hoskison 1983: 47)

(38) nwanwu-kii
chief-DEF
‘the chief’

Koasati has five ‘article suffixes’, which serve to ‘locate a noun in time and indicate that it was previously mentioned’: *-:sáya* ‘the aforesaid’, *-:yólli* ‘the (p.190) very’, *-:ka* ‘the long ago’, *-:kítta* ‘the former’, and *-o:to* ‘the deceased; the long ago’ (Kimball 1991: 405).²⁸

Koasati (Kimball 1991: 407)

(39) á:ti loká:casi:-ka
person orphan-ART
‘the orphaned man’

The precise status of the ‘article suffix’ is not quite clear. Kimball (*ibid.*) observes that four of the five article suffixes are identical to the participial suffixes (*-:sáya* ‘present participle’, *-:yólli* ‘habitual participle’, *-:ka* ‘preterite participle’, and *-:kítta* ‘imperfect participle’). There is no participial suffix corresponding to *-o:to*, nor is there an article corresponding to the future participle.

In the singular the Nasioi definite suffix also indicates noun class and it is reported to have originated from the equivalent of ‘one’.²⁹

Nasioi (Rausch 1912: 113–14)

(40) nanin-u
man-DEF
‘the man’

In the plural only two suffixes (‘bestimmte Pluralartikel’ [definite plural articles]) are employed: *-ni* and *-nanka* (*ibid.*).

The following examples are from the languages in which definiteness is expressed in the form of a free element:

Dutch

(41) de jongen

DEF boy

'the boy'

Guaraní (Gregores and Suárez 1967: 147)

(42) *lo mitã*

DEF boy

'the boy(s)'

(p.191) Hungarian (Moravcsik 1997: 310)

(43) *a fiú*

DEF boy

'the boy'

Pipil (Campbell 1985: 57)

(44) *ne ta:ka-t*

DEF man-ABS

'the/that man'

The translation of the Pipil example is another illustration of the fact that it is often rather difficult to draw a hard and fast line between demonstratives and definite articles. Since Pipil *ne* still has some demonstrative force, it might equally well have been grouped together with the languages in the previous section.

6.2.2.1.3. Numeral classifiers

Numeral classifiers are sometimes employed to express definiteness. Indeed, it has been argued that the numeral classifier is one of the sources from which definite articles can develop (Greenberg 1972: 6), which might have something to do with the anaphoric function of numeral classifiers (see section 5.2.2.1.3). Thus, in Hmong each regular classifier has five forms, which are also said to express (in)definiteness and to give an indication of the size and/or appearance of the referent (section 6.2.1.4) and in Vietnamese the classifier + noun construction is used for anaphoric reference (at least when the classifier bears the main stress):³⁰

Vietnamese (Löbel 1999: 310)

(45) *con qua*

CLF raven

'the (aforementioned) raven'

6.2.2.1.4. Appositional marking of definiteness

There are even languages in which definiteness is expressed by an appositional element. This is perhaps most clearly exemplified in Ngiyambaa, which is not in the sample. In this language definiteness is expressed by a third person absolutive pronoun, which attaches as a clitic to the first word or constituent of the sentence ‘in exactly the same way as when it serves as an anaphoric pronoun ... The nominal(s) in apposition to it may be in any subsequent position in the sentence’ (Donaldson 1980: 128; see also p. 138). For nominals in the ergative case definiteness cannot be expressed in this fashion. Notice that in the example ‘number’ is only expressed on the clitic pronoun in the absolutive.

(p.192) Ngiyambaa (Donaldson 1980: 128)

(46) miri-gu = naŋ-gal	bura:y	gadhiyi
dog-ERG	= 3ABS-PL	child:ABS bite:PAST
'The/a/some dog(s) bit the children'		

6.2.2.2. The expression of indefiniteness

Table 6.2 shows that in most cases indefiniteness (if expressed at all overtly in the NP) is indicated by a reduced, derived, or otherwise related form of the numeral ‘one’. As in the case of definiteness marking, however, it is often difficult to decide whether one is always dealing with the element one is actually looking for. Burushaski, for example, has a certain suffix (*-an*, probably related to *han* ‘one’, with which it may co-occur) which is generally used ‘to isolate a single unit, laying stress on its individuality. It corresponds in force to the English “indefinite article.” Its use is not obligatory’ (Lorimer 1935–8: i.47). However, the same suffix may appear in many other environments, e.g. on nouns that are in construction with a demonstrative adjective, or with predicate adjectives and phrases headed by an adjective (cf. also Berger 1998: 39).

Lango, to take another example, seems to lack the equivalent of an indefinite article. That is, most indefinite NPs are not marked as such, but subject NPs can receive an indefinite interpretation when they are marked with *-mɔ́rɔ̄*.

Lango (Noonan 1992: 162)

(47) twòllórɔ̄	òkàò	átìn
snake:INDEF	3SG:bite:PERF	child
'a snake bit the child'		

Apparently this suffix is almost always found with subjects in existential/locative constructions:

Lango (Noonan 1992: 162)

(48) púnó'mórɔ̄	tíē	ì	ŋè	òt
pig:INDEF	3SG:be:PRES.HAB	in	back	house

'There's a pig behind the house'

Even if a language is explicitly stated as employing an 'exponent of indefiniteness' (Hewitt 1979: 153–4), the element in question may turn up in some rather unexpected environments, which can make its classification as a real indefiniteness marker rather problematic. Consider, for example, this Abkhaz phrase, which can apparently be definite or indefinite (the element *-k* 'INDEF/one' is affixed to the last element of the NP):

Abkhaz (Hewitt 1979: 237)

(49) pš-la-k'
four-dog-one
'(the) four dogs'

(p.193) 6.2.2.2.1. Doubling of indefinite articles

In at least one language in the sample, Tamil, an indefiniteness marker may appear on either side of the head noun. The attributive form *oru* 'a, one' precedes the noun.

Tamil (Asher 1982: 139)

(50) oru raajaa
one king
'a king'

But it is also possible to use the 'nominal form' *oɳṇu* after the noun to indicate indefiniteness:

Tamil (Asher 1982: 140)

(51) kaaru oɳṇu
car one
'a car'

6.2.2.2.2. Indefinite gender

Nama Hottentot has a suffix that expresses what Hagman (1974: 46 ff.) calls 'indefinite gender'. The predicate 'indefinite' is due to the fact that it is 'added to the noun when the speaker does not know what individual person or thing the referent of the noun denotes' (Hagman 1974: 76); the predicate 'gender' is a consequence of the fact that 'no noun stem, be it animate or inanimate or any other category ..., is assigned the indefinite gender in the lexicon. At the same time, any noun stem may be used with this gender, which completely replaces any gender which may be assigned to the noun in the lexicon, conveying the meaning that the referent of the noun stem is unknown, hypothetical, or even non-existent. A Nama noun with indefinite gender is difficult to

translate into English' (ibid. 46). Hagman then goes on to say that the simplest approximation would be to translate it with the English 'indefinite article', but he immediately adds that 'the meanings of the English articles and the Nama genders do not correspond very closely at all' (ibid. 47). According to Greenberg (1978a: 79), however, quoting Meinhof (1909: 48), we are probably dealing here with an (erstwhile) indefinite article that is gradually becoming a neuter marker in a system that already has a masculine and a feminine gender (see Chapter 3, note 22). He mentions two other languages in which a parallel development is supposed to (have) take(n) place: Chinook (Amerind) and Khasi (Austroasiatic).

Alamblak also has a so-called 'indefinite gender', but it seems to function differently (Bruce 1984: 98).

Given that the gender system is regular and obligatorily a part of an NP in third singular forms, conflicts are bound to arise in situations in which the speaker is either unable or unwilling to indicate the gender of an object. In those circumstances the third-person plural marker is employed as an indefinite gender marker. For example, the plural marker (p.194) is used with *yēn* 'child' ... not to indicate the plural number, but to avoid specifying the unknown sex of the child.

Alamblak (Bruce 1984: 98)

(52)	yēn-m	heawrahtm	indom	yamtn
	child-3PL	she_will_bear_them	another	month_in
'She will bear a child in another month'				

Although in Alamblak, as in Nama, NPs end in an element indicating the person, number, and gender of the referent (a so-called PNG-marker; section 3.3.2.2.3), in this language the indefinite gender marker does not originate from an indefinite article but from a third person plural pronoun.

6.3. Possessive modifiers, adpositional modifiers, and relative clauses: the lexical expression of the notion Location in the NP

Lexical expressions of the notion 'location' in the NP can be realized through a variety of constructions, here subsumed under the label 'localizing satellites', of which the possessive modifier and the relative clause are probably most popular across languages.³¹ Possessive modifiers can be divided into (i) possessor pronouns (e.g. 'my' in 'my book'), as indicated in the first column in Table 6.3, and (ii) possessor NPs (like 'the boy' in 'the boy's book'), as indicated in the second column in the same table. Although possessor pronouns are grammatical rather than lexical expressions of the notion location, they are treated here because in many languages they must be used in connection with attributive possessor NPs (see section 6.3.2.1). We will see below that sometimes possessor NPs and relative clauses must be regarded as appositional phrases, so that strictly speaking we cannot say that they are actually functioning as satellites *in* the integral NP. Some grammars contained references to localizing satellites other than those mentioned so far (indicated by * in Table 6.3), such as NPs with the semantic function Source (like 'the train *from Amsterdam*') or Location (as in 'the vase *on the table*'). This is indicated in the fourth column in Table 6.3 under 'other'

(g = adnominal possessor pronoun, E = position of emphatic variant of adnominal possessor pronoun; G = adnominal possessor NP, XR = always cross-referencing of possessor, xr = restricted cross-referencing of possessor; Rel = relative clauses, HI = head-internal relative clause). (p.195)

(p.196)

Table 6.3. Localizing satellites in the NP

Language	g	G	Rel	Other
Abkhaz	E g- N	G N XR	Rel N	* N *
Alamblak	g N	G N	Rel N Rel	
Babungo	E N g	N G	N Rel	N *
Bambara	g N	G N	corel	
Basque	g N	G N	Rel N Rel	* N
Berbice Dutch Creole	g N g	G N G xr	N Rel	
Bukiyip	g N g	G N G xr	Rel N	
Burmese	g N	G N	Rel N	
Burushaski	E g(-) N	G N	Rel N	
Cayuga				
Chinese, Mandarin	g N	G N	Rel N	
Chukchi ¹	g- N	G- N	?	
Dutch	g N g	G N G xr	Rel N Rel	N *
Galela	g N	G N XR	N Rel	
Georgian	g N	G N G	N Rel	
Guaraní	g- N	G N	N Rel	
Gude	g N -g	N G	N Rel	N *
Hittite	g N -g	G N xr	—	
Hixkaryana	g- N	G N XR	—	
Hmong Njua	g N	G N	N Rel	
Hungarian	E N -g	G N G XR	Rel N Rel	
Hurrian	N -g	G N	?	
Ika	g N g	G N G	N Rel & HI	
Kayardild	g N	G N	—	
Ket	g- N	G N XR	N Rel	
Kisi	N g	N G	N Rel	
Koasati	g- N	G N XR	—	
Korean	g N	G N	Rel N Rel	
Krongo	N g	N G	N Rel	N *
Lango	N -g	N G	N Rel	

Nahali	g N	G N	?	
Nama	g N	G N	Rel N	
Nasioi	g N	G N xr	?	
Ngalakan	N -g	G N G xr	—	
Ngiti	g N -g	G N	Rel N	* N
Nivkh	g- N	G N	Rel N	
Nung	N g	N G	N Rel	N *
Nunggubuyu	g N g	G N G	—	
Oromo	Ng	G N G xr	Rel N Rel	
Pipil	g- N -g?	N G XR	N Rel	
Quechua, Imbabura	g N	G N	Rel N & HI	* N
Samoan	(-)g N g	N G	N Rel	
Sarcee	g- N	G N XR	N Rel	
Sumerian	N-g	G N G xr	N Rel	
Tamil	g N	G N	Rel N Rel	
Tsou	N -g	N G	Rel N	
Turkish	E N -g	G N G XR	Rel N	
Vietnamese	N g	N G	N Rel	
Wambon	g- N	G N	Rel N & HI	
West Greenlandic	E N -g	G N XR	N Rel	* N

(1) In his discussion of incorporation in Chukchi, Comrie (1981a: 251) writes that genitives can be incorporated into the head noun (see also Skorik 1961: 392–4; Kämpfe and Volodin 1995: 89).

6.3.1. Attributive possessor pronouns

This section briefly discusses two aspects of attributive possessor pronouns, both of which relate to the formal expression of this noun modifier, namely:³²

- (i) the factor(s) determining the expression of the possessor pronoun in languages that have both bound or free pronominal forms;
- (ii) the factor(s) determining the position of the possessor pronoun relative to the head noun in languages in which these pronouns may occur on either side.

These topics will be treated in turn below in separate subsections, but due to the fact that they are not mutually exclusive, there is a considerable degree of overlap. Several other formal properties of possessive constructions have been ignored, mainly because of the fact that this study focuses on simple NPs.³³

(p.197) 6.3.1.1. Free versus bound forms

The sample contains several languages in which possessor pronouns are expressed both as free (*g*) and as bound forms (-*g/g-*). Abkhaz, Burushaski, Gude, Hittite, Hungarian,

Ngiti, Samoan, Turkish, and West Greenlandic.

Burushaski uses possessive prefixes for what are considered to be ‘inherently possessed’ entities in that language, i.e. entities which always require reference to a possessing entity, such as body-parts and kinship relations (Lorimer 1935–8: i.133, 150; cf. Berger 1998: 79–80). In addition there may appear a free form, presumably for emphatic purposes, yielding a cross-referencing relation between the free pronoun and the prefixed element. For instance:

Burushaski (Lorimer 1935–8: i.133)

(53) (mimo) mi-u
(1PL) 1PL-sons
‘our sons’

Abkhaz, Hungarian, Turkish, and West Greenlandic (Fortescue 1984: 109, 258) normally employ affixes, but in these languages, too, a free prenominal form can be used for emphasis or for stylistic reasons (indicated by E in Table 6.3).

Abkhaz (Hewitt 1979: 116)

(54) (sarà) sə-y°nè
(I) 1SG-house
‘my house’

Hungarian (de Groot and Limburg 1986: 8)

(55) a(z én) kabát-om
DEF (I) coat-1SG
‘my coat’

Turkish (Lewis 1967: 69)

(56) (bizim) ev-imiz
we:GEN house-1PL:POS.SX
‘our house’

The situation in Turkish is more complicated in that ‘pronouns of the first and second persons in the genitive are also used informally as attributive adjectives, i.e. replacing the personal suffixes: ‘our house’ can be *bizim ev*; ‘your street’ *sizin sokak*. The genitive of the third-person pronouns cannot, however, replace the personal suffixes in standard Turkish: *onun ad* instead of *ad- I* [name-3SG:POS.SX] or *onun ad-I* [he:GEN name:3SG:POS.SX] is a provincialism’ (Lewis 1967: 69).

In Gude either a suffix or a free form is used, but the latter only rarely so and then only to express ‘non-inherent possession’ (see below; Hoskison 1983: 44). (p.198) The free variant is actually a pronominal suffix that is attached to ‘an unidentifiable stem beginning in *n*-’ (*ibid.*).³⁴

Gude (Hoskison 1983: 39, 44)

(57) *laa-kii*

cow-3SG

‘his cow’

(58) *naa-kii kuva*

NAA-3SG hut

‘his hut’

Ngiti has different sets of pronouns to express alienable and inalienable possession. The alienable possessive pronouns precede the head noun, the inalienable forms follow the modified noun.

Ngiti (Kutsch Lojenga 1994: 202, 206)

(59) *pbàka mbàra* (short form: *pbà mbàra*)

my arrow

‘my arrow’ (alienable)

(60) *ctsú-du*

hand(s)-my

‘my hand(s)’ (inalienable)

The long and short forms ‘are used interchangeably, depending on speed of speech and emphasis’ (Kutsch Lojenga 1994: 206).

In Samoan possessive pronouns occur in various constructions. The bound form occurs with a ‘possessive preposition’ (*a* or *o*, glossed as POS below; the different forms more or less correlate with the alienable/inalienable distinction; Mosel and Hovdhaugen 1992: 282 f.):

Samoan (Mosel and Hovdhaugen 1992: 277)

(61) *‘o l-o-u fale*

PRES ART-POS-2SG house

‘your house’

Dual or plural possessive pronouns may occur without a possessive preposition:
Samoan (Mosel and Hovdhaugen 1992: 277)

(62) 'o	le	mātou	fale	
	PRES	ART	1.PL.EXCL	house
'our house'				

(p.199) In the emphatic variant the independent pronoun is followed by the emphatic particle *lava*. Apparently this construction is restricted to second person singular possessors:

Samoan (Mosel and Hovdhaugen 1992: 277)

(63) 'o	l-ā	'oe	lava	tama
	PRES	ART-POS	2SG	EMPH child
'your own child'				

Moreover there is also a postnominal variant in which the possessor construction is introduced by a possessive preposition:

Samoan (Mosel and Hovdhaugen 1992: 277)

(64) 'o	le	fale	o	'oe
	PRES	ART	house	POS 2SG
'your house'				

Hittite, finally, used to have suffixes, which were later replaced by free pre-nominal forms. The bound forms, however, were retained in certain expressions (Friedrich 1974: 64 f.).

6.3.1.2. Doubling of the possessor pronoun

In the following languages possessor pronouns, whether bound or free, are attested on either side of the head noun: Babungo, Berbice Dutch Creole, Bukiyp, Dutch, Gude, Hittite, Hungarian, Ika, Ngiti, Nunggubuyu, Pipil (see note 35), Samoan, Turkish, and West Greenlandic. Several of these languages, however, have already been referred to in the previous section; only Babungo, Bukiyp, Berbice Dutch Creole, Dutch, Ika, and Nunggubuyu remain to be discussed.³⁵

Whereas in Babungo emphasis determines the placement of the attributive possessor pronoun, in Bukiyp the position of the possessive modifier (pronoun or NP) depends on the presence of other modifiers. If the possessive modifier is the only modifier it appears before the noun, but if, for example, a demonstrative, numeral, or adjective precedes the head noun, the possessive modifier 'is permuted to postnuclear position' (Conrad 1991: 57).

In Berbice Dutch Creole the postnominal variant has a relator in the form of *fan* ‘from/of’; it is not clear what determines the choice, but perhaps this is due to the influence of Dutch (see below):³⁶

Berbice Dutch Creole (Kouwenberg 1991: 113)

(65) ʃi	tun
	3POS field
	‘his field’

(p.200)

(66) di	kali	kujar	fan	ɛke
	the	small	canoe	from/of 1SG
	‘my small canoe’			

Dutch possessive pronouns normally precede the noun but in the presence of certain other prenominal modifiers (notably determiners) a postnominal prepositional phrase is used (no longer involving the special possessive form):

Dutch

(67) mijn	boek-en
	my book-PL
	‘my books’

(68) een	boek	van	mij
	a	book	of me
	‘a book of mine’		

The normal position of the possessor pronoun in Ika is before the head noun, but according to Frank (1990: 4, 41 f.) it may also appear in postnominal position.

It is not quite clear what causes possessor pronouns to precede or to follow the head noun in Nunggubuyu. Perhaps this is related to the fact that the element in question is actually an independently referring expression by itself, i.e. has an appositional relationship with the head noun (Heath 1984: 250, 497 ff., 545 ff.).

6.3.2. Attributive possessor NPs

This section focuses on just two aspects of attributive possessor NPs: (i) cross-referencing and (ii) position relative to the head noun. As regards the first aspect, the discussion will include only those languages in which cross-referencing is not a feature

of every NP that contains an attributive possessor NP. As for the position of the possessor NP, I will only consider languages in which this phrase can occur on either side of the head noun.

6.3.2.1. Cross-referencing

In certain languages an attributive possessor NP may or must co-occur with a pronominal element, both constituents referring to the possessing entity. In other words, in these languages there is cross-referencing (= XR/xr) between the attributive possessor NP and some pronominal element. If such a relationship is obligatory, this is indicated by capitals (XR) in Table 6.3. If the extra pronominal element does not always occur (but only e.g. with certain head nouns or possessing entities), the same characters are given in lower case (xr). This section is restricted to languages in which the possessor NP is *not* always cross-referenced by a possessor pronoun (i.e. xr-languages): Berbice Dutch Creole, Bukiyp, (p.201) Dutch, Hittite, Nasioi, Ngalakan, Oromo, and Sumerian. Below the conditions are listed under which this pronoun appears, to the extent that this could be determined on the basis of the available data.

It is not clear when cross-referencing occurs in Berbice Dutch Creole and Bukiyp; compare:

Berbice Dutch Creole (Kouwenberg 1991: 113)

(69) di	j ^e rma	papa
the	woman	father
'the girl's father'		

(70) di	potman	j ⁱ	toro
the	old man	3POS	eye
'the father's eye'			

Bukiyp (Conrad 1991: 60)

(71) Manohweh	buwul
Manohweh	Pig
'Manoweh's pig'	

Bukiyp (Conrad 1991: 59)

(72) nebe-nali	aninú	éman-i-hw	awehw
important-CL7.SG	father	he-POS-CL12.SG	song
'important father's song'			

Instead of the postnominal construction with *van* 'of' (see below), spoken Dutch has a cross-referencing construction (involving a reduced form of the possessive pronoun)

which only occurs with prenominal animate (mostly human) possessors (Geerts et al. 1984: 208–9):

Dutch

- (73) dat meisje dør fiets [dør = reduced form of haar ‘her’]
that girl her bike
'that girl's bike'

The only information Friedrich (1974: 122) gives about the use of a possessor pronoun in combination with a possessor NP in Hittite is that it is often attested in legal documents and in the older texts:

Ein umständlichere Art der Genitivbezeichnung ist die mittels Zufügung des Possessiv-pronomens (‘des Marines sein Kopf’ ...). Sie ist besonders in den Gesetzen (aber in alter Sprache auch sonst) beliebt ...

[An elaborate way to express the genitive is by adding the possessive pronoun (‘the man his head’ ...). This occurs especially in legal texts (but also in the ancient language).]

(p.202) Nasioi possessing entities are only cross-referenced in the case of a possessive genitive (‘possessiver Genitiv’), which includes NPs referring to human possessors, but not those involving proper names, kinship terms, or inanimate nouns:

Nasioi (Rausch 1912: 119, 120)

- (74) nánin bakana danko
man his spear
'the man's spear'

- (75) Máteasi bauran
Mateasi daughter
'Mateasi's daughter'

In Ngalakan the head noun is ‘almost always appropriately suffixed to cross-reference the genitive-marked possessor, though this can be omitted, ... Possessed and possessor, if both expressed by nouns, may occur in either order’ (Merlan 1983: 44; see also p. 73).

A possessor NP may precede or follow the head noun in Oromo, but when it precedes the presence of a third person pronoun is required (as well as a so-called linker clitic = LIN; Stroomer 1987: 174 f., 180). Compare:

Oromo (Stroomer 1987: 178, 180)

- (76) gogaa k'urt'ummii guddiyoo aa

skin fish big LIN = GEN

'the skin of a big fish'

(77) sareeni tana k'eensa íšii aa c'aafuu

dog this paw her LIN = GEN dirt(y)

'the paw of this dog was dirty'

In Sumerian the possessor NP usually follows the head noun:

Sumerian (Thomsen 1984: 91)

(78) é lugal-ak

house king-GEN

'the house of the king'

As in Oromo the possessor NP may also precede (the so-called anticipatory genitive), in which case a pronominal element is required:

Sumerian (Thomsen 1984: 91)

(79) lugal-ak é-ani

king-GEN house-his

'the house of the king'

(p.203) 6.3.2.2. Doubling of the possessor NP

In the following languages the adnominal possessor NP may occur on either side of the head noun: Berbice Dutch Creole, Bukiyp (section 6.3.1.2), Dutch (see above), Georgian, Hungarian, Ika, Ngalakan (see above), Nunggubuyu, Oromo, Sumerian, and Turkish.

As in the case of possessor pronouns, the postnominal variant requires a relator both in Dutch and Berbice Dutch Creole (for other possibilities see above):

Dutch

(80) a. Peter-s boek b. het boek van Peter

Peter-GEN book the book of Peter

'Peter's book' 'Peter's book'

Berbice Dutch Creole (Kouwenberg 1991: 113)

(81) di rai fan di trakta

the noise from/of the tractor

'the noise of the tractor'

In Georgian and Hungarian the possessor NP most commonly precedes the head noun, but it may also appear after the noun. Apart from the fact that in Georgian 'postposed genitives are associated with archaic or solemn style' (Testelec 1997a: 252), possessor NPs are also placed after the head noun so as to reduce the distance between the head of the possessor NP and the head of the matrix NP (see Chapter 9 on the *Head Proximity Principle*). This is exemplified in the following example, in which the head of the possessor NP is modified by a relative clause which would otherwise have occurred in between the head of the possessor NP (*hatred*) and the head of the matrix NP (*reason*):³⁷

Georgian (Testelec 1997a: 252): NG order

(82)	mizez-i	sizulvil-isa,	romel-i-c	didixania
	reason-NOM	hatred-GEN	which-NOM-REL	long
	gv-açux-eb-s			
	1PL.O-trouble-PRS.3SG.S			
'the reason of the hatred which has troubled us for a long time'				

Compare:

Georgian (Testelec 1997a: 252): GN order

(83)	?sizulvil-is,	romel-i-c	didixania
	hatred-GEN	which-NOM-REL	long
	gv-açux-eb-s,		mizez-i
	1PL.O-trouble-PRS.3SG.S		reason-NOM
'the reason of the hatred which has troubled us for a long time'			

(p.204) In Hungarian the prenominal possessor NP may or may not have the case marker, whereas the postnominal one must have it:

Hungarian (Edith Moravcsik, personal communication)

(84)	János(-nak)	a	kalap-ja
	Janos(-GEN)	DEF	hat-3SG
'John's hat'			

(85)	a	kalap-ja	János-nak
	DEF	hat-3SG	Janos-GEN

'John's hat'

The possessor NP in Ika may occur in either side of the head noun, but in the case of ownership ‘the possessor phrase is much more likely to follow the possessed item’ (Frank 1990: 42):³⁸

Ika (Frank 1990: 42)

- (86) peri zΛ-džua
dog GEN-blood
'dog's blood'

- (87) tšinu in ?gui zei
Pig someone GEN
'someone's pig'

The order of genitive and head is not strictly fixed in Nunggubuyu (Heath 1984: 500), which may have something to do with the appositional character of the possessor NP (*ibid.* 505). The factors determining the position of the possessor NP in Ngalakan, Oromo, and Sumerian are not given either (see above; cf. Merlan 1983: 44, 73; Stroomer 1987: 180; and Thomsen 1984: 91).

As a rule a possessor NP precedes the head noun in Turkish, but in colloquial speech it may also be placed after the noun.³⁹ Lewis (1967: 242–3) writes:

In informal speech the answer to a question like ‘what’s that place over there?’ may well be in the form *ev-i çiftçi-nin* [house-3SG:POS.SX farmer-GEN ‘the farmer’s house’—JR]. The inverted order is even more likely if the phrase is part of a longer sentence, e.g. *evi büyük, çiftçi-nin* ‘his house is big, the farmer’s’. This may look as if the qualifier (p.205) *çiftçinin* is added as an afterthought, but in fact this is at least as common a form of sentence in the spoken language as the formal *çiftçi-nin evi büyük*.

6.3.3. Relative clauses

Once again it must be emphasized at the outset that this is not an extensive discussion about a certain group of satellites in the NP, because this study is mainly about simple NPs.⁴⁰ I will only attempt to give some additional information about positional properties or tendencies of the *adnominal* restrictive relative clause in languages in which such clauses are attested on either side of the head noun. This restriction also accounts for the fact that certain types of relative clauses are ignored in the present context, such as the head-internal relative clause, the extraposed relative clause, the appositional relative clause, the adjoined relative clause, and the correlative construction.⁴¹

The head-internal construction, which is employed in Ika (Frank 1990: 5), Imbabura Quechua, and Wambon (de Vries 1989: 110–11), is not included because—as the name

suggests—the head noun is an integral part of the relative clause, which is itself usually a nominalized construction. Consider this example from Imbabura Quechua, in which *jari* ‘man’ (with the semantic function Benefactive) can be regarded as the internal head (Cole 1982: 50, 55 f.):

Imbabura Quechua (Cole 1982: 56)

(88)	Marya	jari-paj	ruwana-ta	rura-shka-ka
	María	man-for	poncho-ACC	make-NLZR-TOP
	Agatu-pi-mi		kawsa-n	
	Agato-in-VAL	live-3		

‘The man for whom María made a poncho lives in Agato’

Extraposed relative clauses are excluded here as well, because they are expressed separately from the (relativized) matrix NP in the sentence.⁴² This kind of discontinuity is usually attributed to heaviness: the tendency for speakers to place ‘heavy’ (i.e. complex) structures toward the end of the sentence (Dik 1997: 410). Extraposition is exemplified in the following example, which is also from Imbabura Quechua. Both the head and the extraposed relative clause must now be case marked:

Imbabura Quechua (Cole 1982: 51)

(89)	kwitsa-ta	juya-ni	Juan-wan	tushu-shka	ka-shka-ta
	girl-ACC	love-1	Juan-with	dance-NLZR	be-NOM-ACC

‘I love the girl who had danced with Juan’

(p.206) Samoan has two types of relative clause. One is an integral part of the NP, the other (the so-called ‘syndetic relative clause’; Mosel and Hovdhaugen 1992: 317) is analyzed as an apposition. Its nucleus is formed by the relative pronoun *ē*:

Samoan (Mosel and Hovdhaugen 1992: 317)

(90)	'O	lea	'o	le	tama	'ita	'i	l-ē
	PRES	that	PRES	ART	woman		ART-REL	
	na	tātou	ō	'i	ai			
	PAST	1.PL.IN	go:PL	LD	ANAPH			

‘She is the woman we went to find’

[lit. ‘That is the woman the one we went to’]

Adjoined relative clauses do not figure in this section for the simple reason that they, too, do not seem to be part of an NP; they either precede or follow the clause rather than the head noun and are formally indistinguishable from adsentential clauses. Adjoined relative clauses are particularly common in Australian languages. This

example is from Ngiyambaa; (i) is a literal translation, (ii) is a more colloquial alternative:

Ngiyambaa (Donaldson 1980: 299)

(91)	yalama-nhi-ba	mayi,	guwayubu = na	yuwa-nha
	tired-PAST-SubSx	person:ABS	still = 3ABS	lie-PRES
(i) 'person was tired, still he is sleeping'				
(ii) 'the person who was tired is still sleeping'				

As this example shows, the adjoined clause can often be interpreted as an adverbial clause. This is also illustrated by these examples from Warlpiri (see also Simpson 1983):⁴³

Warlpiri (Hale 1976: 78–9)

(92)	ŋatjulu-ļu	ɸ-ɳa	yankiri-ļi	pantu-ɳu	kutja-ipa	ŋapa	ɳa-ɳu
	I-ERG	AUX	emu	spear-PAST	Comp-AUX	water	drink-PAST
'I speared the emu which was/while it was drinking water'							

With the adjoined clause preposed, we get:⁴⁴

Warlpiri (Hale 1976: 78–9)

(93)	yankiri-ļi	kutja-ipa	ŋapa	ɳa-ɳu,	ŋatjulu-ļu
	emu-ERG	COMP-AUX	water	drink-PAST	I-ERG
	ɸ-ɳa	pantu-ɳu			
	AUX	spear-PAST			
'The emu which was drinking water, I speared it'					
'While the emu was drinking water, I speared it'					

(p.207) In Nagalakan this kind of 'generalized subordinate clause' (Merlan 1983: 135) usually gets a relative clause interpretation when the modified constituent immediately precedes the subordinate clause, as in:

Ngalakan (Merlan 1983: 138)

(94)	mu-yal̩kič	yimi	bareñ-gen	dar̩?-ga?
	MU-dilly_bag	1IN	DU/MU-hang_up-PP-SubSx	tree-LOC
	ɸ-ɳɔr̩?-miň	we?	-ga?	
	3SG-fall-PP	water-LOC		
'The dilly bag we hung on the tree fell into the water'				

The correlative construction (see Comrie 1981b: 139–40; note that Keenan (1985: 163) uses ‘corelative’) is generally characterized by the fact that it contains an NP with a distinctive element, the correlative marker (COR), which is referred to anaphorically by an element in the main clause (Comrie 1981b: 139–40). This example is from Hindi, but the correlative construction was also used in e.g. Hittite (Friedrich 1974: 167–8):⁴⁵

Hindi (Keenan 1985: 164)

(95)	Jis	a:dmi	ka	kutta	bema:ar	hai,	us	a:dmi	ko
	COR	man	GEN	dog	sick	is,	that	man	DO
	mai	ne		dekha					
	I		ERG	saw					
‘I saw the man whose dog is sick’									
(lit. ‘Which man’s dog was sick, that man I saw’)									

Since Bambara (Kastenholz 1989: 199), Hittite, Ngalakan, and Kayardild (Evans 1995: 517) do not employ adnominal relative clauses, they have no entry in the third column of Table 6.3. The evidence in relation to Nunggubuyu is not conclusive. It is argued that relative clauses in this language are ‘best viewed as separate, appositional units’, which ‘need not be directly adjacent to the head noun (which indeed may be merely covert); instead they are often separated from it by pauses and/or other intervening constituents’ (Heath 1984: 500). Although they usually follow the head noun, they may also follow (or precede) the main clause (*ibid.* 570).

Hixkaryana (Derbyshire 1979: 26, 132) and Koasati (Kimball 1991: 525–7) have no entry either, because they are said to lack relative clauses altogether and functionally equivalent constructions do not seem to involve attributive pre- or postnominal phrases.

(p.208) 6.3.3.1. Doubling of the relative clause

Taking into consideration the fact that certain types of relative clauses have been ignored (see above), doubling of the relative clause is a feature of at least the following languages in the sample: Alambalak, Basque, Dutch, Georgian, Hungarian, Korean, Nunggubuyu (but see above), Oromo, Tsou, and Tamil (see Hawkins 1983: 339).⁴⁶

The position of the relative clause in Alambalak mainly depends on its size, i.e. short clauses (which are attested most often) usually precede the head noun whereas long ones tend to follow the head noun. Furthermore, short relative clauses commonly occur without a so-called relativizer, i.e. a demonstrative element that is incorporated into the verb of the relative construction:

Alambalak (Bruce 1984: 108, 109)

(96)	ind	habhi	kmi	na	ind-kfämë-t
	DEM	small	place	I	DEM-said-3SG.F
‘the small place about which I spoke’					

(97) n̩em Ukarumpakorn yirn̩e w̩ik-r
we Ukarumpa.to went week-3SG.M
'the week we went to Ukarumpa'

It is not explained what determines the position of the relative construction in Basque (Saltarelli 1988: 36 ff.), but the postnominal construction seems to be the marked alternative.

Dutch, Georgian (Testelec 1997a: 253; Harris 1994), and Hungarian have a non-finite (participial) construction that precedes the noun as well as a finite clause in postnominal position.

Dutch

(98) de in de tuin zittende man
the in the garden sitting man
'the man (who is/was) sitting in the garden'

(99) de man die ik gisteren ontmoet heb
the man who I yesterday met have
'the man I met yesterday'

At least in Dutch the participial construction is more like a qualifying modifier, whereas the finite construction rather serves as a localizing satellite in that it (p.209) helps the addressee to identify the referent of the matrix NP (here *the man*; cf. Rijkhoff 1997: 362; Siewierska et al. 1997: 786).

In Korean relative clauses normally precede the head noun; only rarely do they follow:

Korean (Lee 1989: 174)

(100) n̩ega san c̩eg
I bought book
'the book which I bought'

(101) c̩eg n̩ega san
book I bought
'the book which I bought'

In Oromo a relative clause normally follows the noun, but may precede it 'for purposes of foregrounding, or for avoiding ambiguous or cumbersome sequences of post-nominal

modifiers' (Gragg 1972: 164; see also Gragg 1976: 191).

Oromo (Stroomer 1995: 117)⁴⁷

(102)	ani	intala	kalee	d'ufteeti	
	ani	intala	kalee	d'uf-te	ee-ti
	1SG:S	yesterday	come-3SG.F.PAST.SubV	LIN-SEP	
	hime				
	him-e				
	tell-1SG.PAST				
	'I told the girl who came yesterday'				

Finally, in Tsou a relative clause only follows the noun as an afterthought (Szakos 1994: 162).

6.3.4. Other kinds of localizing satellites in the NP

Relatively few grammars contain information regarding other kinds of localizing satellites in the NP. In the sample these are the grammars of Abkhaz, Babungo, Basque, Dutch, Gude, Krongo, Nama, Ngiti, Nung, Imbabura Quechua, and West Greenlandic.

In Abkhaz, for instance, the 'ablative' postpositional phrase may precede or follow the head noun:

Abkhaz (Hewitt 1979: 58)

(103)	Očamčèra-n	+t ^o	à-jyab	
	Očamčira-from	ART-girl		
	'the girl from Očamčira'			

(p.210)

(104)	à-jyab	Očamčèra-n	+t ^o	
	ART-girl	Očamčira-from		
	'the girl from Očamčira'			

There appear to be differences as regards pre- or postnominal position depending on the semantic function of the satellite NP in question, e.g. an attributive postpositional phrase with the function Benefactive ('the presenter *for* Mary') most often appears before the noun. Not all semantic functions can be expressed this way. For instance, Abkhaz has no structural parallel to 'the boy *with* the girl'; instead 'with the girl' must be realized as a relative clause (Hewitt 1979: 115).

In Babungo only prepositional phrases indicating a location can modify an NP (Schaub 1985: 75), whereas in Basque NPs with a variety of semantic functions (e.g. Benefactive, Comitative, Locative) may function as localizing satellites in the NP, provided they carry the suffix *-ko*, which is called a relativizing case suffix (RC), but which ‘strictly functions as a structural particle in this context and can be used to mark various classes of postpositional constructions’ (Saltarelli 1988: 197). The following rather extreme example illustrates:⁴⁸

Basque (Saltarelli 1988: 80)

(105) mendi-ko	baserri-ko	amona	zaharr-aren
mountain:LOC-RC	farm:LOC-RC	grandmother	old-SG.GEN
aurpegi-ko	azal-aren	zimurr-ak	
face:LOC-RC	skin-SG.GEN	wrinkles-PL.ABS	

‘the wrinkles of the skin on the face of the old grandmother in the farm of the mountains’

In Dutch, Gude, and Krongo such localizing satellites all occur in postnominal position.

Dutch

(106) de man in de tuin
the man in the garden
‘the man in the garden’

Gude (Hoskison 1983: 63)

(107) maarə a gəra
oil at pot
‘the oil in the pot’

Krongo (Reh 1985: 149)

(108) káaw nk-úbbà
person ABL-far away [my glosses/translation]
‘a person from far-away/a stranger’

(p.211) Although the following example is taken from a section on nominalized clauses in Nung (Saul and Freiberger Wilson 1980: 15–16), there do not seem to be any overt signs that we are dealing with a (subordinate) clause here:⁴⁹

Nung (Saul and Freiberger Wilson 1980: 16)

(109) cáh tú pá dù cháng áhn hñ̄n nī, muñhñ vñ̄

CLF	CLF	fish	at inside	CLF	house	FOC	he	say
'the fish inside the house said ...'								

Localizing satellites precede the head noun in Imbabura Quechua: Imbabura Quechua (Cole 1982: 77)

(110) Kitu-manda	runa
Quito-from	man
'the man from Quito'	

This is an example of a ‘simple postposition’: a single morpheme suffixed to a noun. Quechua also has ‘complex postpositions’, which consist of a noun stem followed by a postpositional suffix.⁵⁰

Imbabura Quechua (Cole 1982: 75)

(111) wasi	uku-pi/man/manda
house	interior-in/to/from
'inside/to/from the house'	

The last example illustrates that in West Greenlandic case marked NPs in adverbial function can also occur as modifiers in the NP:

West Greenlandic (Fortescue 1984: 112)

(112) qali-a-ni	sinittarvik
loft-its-LOC	bedroom
'a bedroom in the (house's) loft'	

6.4. Conclusion

We have now concluded the presentation of NP-internal constituents. To recapitulate briefly, the discussion began in Chapter 2 with a new typology of nominal subcategories (*Seinsarteri*) and in Chapter 3 I showed how different forms of noun classification influence the form and order of constituents both inside and outside the domain of the NP. Chapter 4 focused on noun modifiers which further (p.212) specify more or less inherent properties of the referent as it is defined by the head noun: *qualifying operators* (nominal aspect markers) and *qualifying satellites* (here I restricted myself to one particular kind, namely attributive adjectives). Chapter 5 was concerned with *quantifying operators* and *quantifying satellites*, i.e. grammatical and lexical elements which pertain to quantitative properties of the referent, and the present chapter has dealt with modifier categories in the NP which relate to locative properties of the referent, *localizing operators* and *localizing satellites*.

Having established some of the most important properties of (first order) nouns and their modifiers, we can now proceed to consider how NP structure is to be captured in a general theory of grammar, in particular Dik's *Functional Grammar* (Dik 1997). I will be concerned chiefly with two issues, namely how the various operators and satellites can be represented in a semantic model of the noun phrase (Chapter 7) and which principles determine their position in a linguistic expression (Chapters 8–10).

Notes:

(1) Recall that adjectives may also function as localizing satellites (as in 'No, I want the RED apple', where the adjective is used contrastively, serving to identify the referent). However, since attributive adjectives mostly seem to be used as qualifying satellites, they were treated in Chapter 4. Relative clauses and embedded NPs can also be used as qualifying satellites, especially in languages without adjectives, but since they are more often used as localizing satellites in languages that have a distinct category of adjectives they are treated here.

(2) Since articles are generally regarded as markers of 'weak deixis' (section 6.2.2), I treat them here in the context of localizing operators, but I will argue in Chapter 7 that they are better characterized as discourse operators, because they relate to the referent as a discourse entity.

(3) The term 'locative modifier' covers various phrases whose semantic function relates to a location in the spatial dimension, such as 'in the park', 'under the house', 'through the forest', 'over the table', 'with the dog', 'behind the garage', 'next to her', 'against the wall', 'by the river', 'at the post office', 'opposite the church', etc.

(4) Lyons (1977: chapter 15.2.); Greenberg (1985); Anderson and Keenan (1985: 277 ff.); Dik (1997: 180f.); Gildea (1993); Himmelmann (1997).

(5) Cf. Ultan (1978a) Sometimes the origin of the demonstrative itself can be traced back to a lexical source (but cf. Diessel 1999a: 44). In the case of Tuscarora, for example, Mithun Williams (1976: 33) writes that the deictic *kyé:nv*: 'this' derives from *k + yenv*: 'I am holding it'. Copeland (1998) has argued that in Tarahumara the form *ma* (originally 'hand') now 'functions as a deictic to identify shifting spatial referents'.

(6) See also, for example, the following two statements by Heine: 'Determiners in a narrow sense ... include demonstrative and possessive adjectives. Their major "function" appears to be deictic: they relate the concept expressed by the head noun to the relevant speech context with regard to participants ("my", "your", etc.), place ("this", "that", etc.) and time' (Heine 1980: 185). 'Locative constructions are probably the most common source for expressing possession in African languages. Underlying this transfer from the spatial domain to the domain of possession, there appears to be an implication of the following kind: what is at my place belongs to me' (Heine 1990: 144).

(7) Recall that relative clauses may also function as qualifying satellites (see note 1 and Chapter 4).

(8) Rijkhoff (1989); Dik (1997: 129 f.); see also Jakobson (1971b) on shifters.

(9) Anchoring is also one of the ways to ‘ground’ an NP. ‘Grounding is the primary way in which speakers make an NP relevant. To ground a noun phrase is to locate its referent in conversational space by relating it to a referent whose relevance is clear, that is, to a Given referent in the immediate context’ (Fox and Thompson 1990: 300; see also e.g. Hannay 1985). Currently much research on such constructions is being done in discourse representation theory (DRT) in the larger context of *presupposition* (Kamp and Reyle 1993; van der Sandt 1992).

(10) There are interesting relations between different kinds of localizing operators and satellites. For example, Perkins (1992) has suggested that there is an inverse correlation between the number of distinctions in demonstratives (*localizing operators*) and the number of relativizable positions in Keenan and Comrie’s *Relativization Hierarchy* (Keenan and Comrie 1977).

(11) See also Foley (1986: 201): ‘Subordinate clauses in Papuan languages always function to background given information, and correspond to two different constructions in more familiar languages, adverbial clauses and relative clauses. Both of these constructions express background information in a sentence: adverbial clauses provide a temporal setting or background condition or cause for the asserted new information in the main clause, e.g. *while I sat under the tree, an apple hit me on the head* or *if he comes, I will leave*; while relative clauses provide crucial background information for the identification of the referents of the head nouns: *the pig which has white spots* or *the pig which I killed*. The close relationship between these two types of subordinate clauses is readily apparent in many Papuan languages, in which they are formally very similar or even identical.’

(12) See also Saltarelli (1988: 80) on Basque; Lee (1989: 120) on Korean; Rausch (1912: 120) on Nasioi; Thomsen (1984: 54) on Sumerian; Asher (1982: 63) on Tamil; Fortescue (1984: 144) on West Greenlandic.

(13) Traugott (1975); Heine (1990); Traugott and Heine (1991).

(14) For detailed discussions of these and other aspects of (spatial) deixis the reader is referred to Lyons (1977: chapter 15); Levinson (1983: chapter 2); Rauh (1983); Anderson and Keenan (1985: 259–308); Greenberg (1985: 271–87); Dik (1997: 180–3).

(15) Cf. Greenberg (1978a, 1985); Noonan (1985); Gildea (1993); Frajzyngier (1991, 1997); Diessel (1999a,b).

(16) See also Moravcsik (1997: 319), who argues that the adnominal demonstrative in Hungarian arose ‘from a pronominal demonstrative to which a noun phrase was appositionally adjoined. Thus, the origin of an expression like *ez a barna kalap* “the brown hat” was *ez, a barna kalap* “this, the brown hat”. A little further she writes (*ibid.*): ‘The adnominal demonstrative in French seems to have run the same historical course (cf. McCool 1984, 1993). Lehmann (1995: 37–39) has suggested that this is in fact the cross-linguistic norm (cf. also Ultan 1978a, Greenberg 1978a).’ Cf. also Lehmann (1984: 177) on apposition in Proto Indo-European.

(17) The demonstrative suffix in Gude may appear on the noun or on the numeral: (Hoskison 1983: 51): ‘Normally a modifying cardinal number follows the head noun and its other modifiers. However, the cardinal number may precede the demonstrative suffix ... in which case the suffix attaches to the number.’ Compare:

(1) a.	fəzə-tsə	mak	b.	fəzə	makə-tsa
	year-DEM	three		year	three-DEM
	'those three years'			'those three years'	

(18) In the following languages, too, demonstratives ‘often co-occur with a definiteness marker’ (Moravcsik 1969: 76): Efate (Austric), Samoan (Austric), Santo (Austric), Tanna (Austric), Macedonian (Indo-Hittite), Maasai (Nilo-Saharan), Luganda (Niger-Kordofanian), Loma (Niger-Kordofanian), Lyele (Niger-Kordofanian), Bengali (Indo-Hittite), Athabascan (Na-Dene), Egyptian Arabic (Afro-Asiatic), Irish (Indo-Hittite), and Ancient Greek (Indo-Hittite). Note that in some languages possessor pronouns may also occur with a ‘definite article’ (e.g. Italian, Bulgarian; see e.g. Manzelli (1990) about this phenomenon in the European languages). Cf. also Plank (1991: 6): ‘If articles co-occur with possessive determiners, then demonstratives do so, too.’

(19) But at least in Berbice Dutch Creole and Hungarian the article only occurs in definite NPs.

(20) Greenberg (1972, 1978a: 78) has pointed out that if the numeral classifier occurs with other constituents, it will first do so with demonstrative pronouns, which is perhaps not very surprising in view of the anaphoric character of numeral classifiers.

(21) For general overviews of the way (in)definiteness is expressed cross-linguistically I refer to Moravcsik (1969), Krámský (1972), Dryer (1989a), Heine (1997: 66–82), and Lyons (1999). On the use of articles in adpositional phrases, see Himmelmann (1998). Notice that in his article on article- noun order, Dryer (1989a) treats various elements as articles which have been left out of consideration here (e.g. noun (phrase) markers as they are attested in Samoan and Fijian); see also section 9.3.3.4.1.

(22) To give another example, in Ossetic (Indo-Hittite) definiteness was indicated by shifting the stress to the second syllable. This is probably due to the incorporation of an erstwhile definite article (Abaev 1964: 12; see also Chapter 3, note 19). Schroeder (1999: 40–5) offers a detailed overview of the ways in which definiteness is coded in the Turkish language.

(23) In many languages the third person possessive suffix is used to convey the notion of definiteness (cf. Menges 1968: 113 on the Turkic languages; Benzing 1985: 12 on Kalmyk; Comrie 1988: 465 on this phenomenon in the Uralic languages; see also Siewierska et al. 1997: nn. 14, 32).

(24) The same kind of syntactic variation to express definiteness and indefiniteness was used in Old Georgian (Testelec 1997a: 247): ‘The indefinite article was *erti* which functioned as an article only if postposed, and as a numeral “one”, if preposed. The demonstrative pronouns *igi*, *ese*, *ege* “this, that” were preposed; if postposed, they functioned as definite articles.’

(25) This was more or less explicitly stated to be the case for the following languages in the sample: Alambalak (Bruce 1984, 100), Babungo (Schaub 1985: 73, 97, 192), Mandarin Chinese (Li and Thompson 1989: 86, 131), Korean (Kim 1987: 895), Nama

Hottentot (Hagman 1974: 74), Oromo (Stroomer 1987: 88, 1995: 48), and Tamil (Asher 1982: 61, 139 f.).

(26) The only reference to a (possible) definiteness marker in Chukchi that I found reads as follows (Bogoras 1922: 716): '[A certain suffix] seems to express an emphatic form. Sometimes it corresponds to the definite article or designates an object as referred to before.'

(27) See e.g. Merlan (1982: 44, 93) on Mangarayi *gi-* and the functions of case and ‘number’ marking in relation to (in)definiteness and referentiality; and Heath (1984: 169–72) on the function of the continuous prefix in Nunggubuyu.

(28) Cf. Boas (1911a: 39): ‘Tense classes of nouns are not rare in American languages. As we may speak of “a future husband” or of “our late friend”, thus many Indian languages express in every noun its existence in present, past, or future, which they require as much for clearness of expression as we require the distribution of singular and plural.’ See also Chapter 1 on languages with a parts-of-speech system of type 5.

(29) ‘Die Substantive zerfalien, wie bei den Zahlwörtern zu ersehen ist, in Wertklassen. Als Wertzeichen dienen im Singular gewisse Suffixe. Diese Suffixe geben die Art und Beschaffenheit des Gegenstandes an und fungieren zu gleicher Zeit als bestimmter Singularartikel. Das einer jeden Klasse zukommende Suffix ist an dem Zahlwort “eins” erkenntlich, mit welchem es auch in den meisten Fällen buchstäblich übereinstimmt’ (Rausch 1912: 113–14). [As can be seen in the numerals, the substantives divide into semantic classes. This is indicated in the singular by certain suffixes. These suffixes indicate the kind and the quality of the object and at the same time serve as definite singular articles. In a suffix of any class one can recognize the numeral ‘one’, with which it is also formally identical in most cases.]

(30) Cf. also Sun (1988) on the discourse function of numeral classifiers in Mandarin Chinese.

(31) Besides sharing the same function in the NP, the possessive modifier and the relative clause are also similar in other respects, such as morphological marking. In many languages there is an obvious formal relation between genitive and relative marking, which is usually thought to have developed from an erstwhile pronoun (see e.g. Aristar 1991: 13).

(32) I am aware that in certain languages there is no separate category of possessive pronouns; instead personal pronouns may be used (according to Ultan (1978b: 36), who used a random sample of seventy-five languages, ‘most possessive pronouns or adjectives are formally identical or similar to object pronouns’). I use the phrase ‘attributive possessor pronoun’ for any pronominal element in the NP denoting some possessing entity. In turn, the phrase ‘possessing entity’ is used rather liberally here, since it includes human, non-human animate, and inanimate entities. Usually only humans are seen as true possessors.

(33) For instance, in addition to pronominal possessor prefixes Hixkaryana has possessive suffixes, which also mark ‘number’ and present/past distinctions (Derbyshire 1979: 69 f., 96 f.). These and other aspects (genitive marking, de possession, etc.) are generally ignored here, but see e.g. Ultan (1978b); Limburg (1985); Seiler (1983). For another typological overview see Croft (1990a: 26–39); for an overview of

various properties of attributive possessor pronouns in all European languages, see Manzelli (1990).

(34) Recall that the suffix *-kii* is also used to mark definiteness in Gude (section 6.2.2.1.2). As to the independent form, Hoskison (1983: 44) adds: ‘In other Chadic languages the independent possessive pronominal is often formed by adding a possessive suffix to the word for “head” (Note Hausa: *kai* “head”, *kaina* “my head, myself”). This practice may be the origin of the Gude forms as well, but synchronically the Gude independent possessive pronominals are distinct from inflected forms of *na* “head”.’

(35) Pipil also has ‘possessive suffixes’, *-yu* and *-w(an)*, which occur with e.g. body-parts and kinship relations (Campbell 1985: 42–6); it is not clear which of these suffixes are pronominal elements.

(36) Only for the third person singular is a distinct possessive pronoun available in Berbice Dutch Creole, *si*, which is in free variation with the general form *ori*.

(37) Fähnrich (1986: 156) writes that the possessor NP most frequently precedes the noun in Georgian and that postnominal placement requires the ‘long form’ of the possessive modifier (i.e. with the *-a* suffix).

(38) The genitive marker has two forms in Ika: *zei* and *zϕ-*. ‘When a genitive precedes its head noun and the head noun is one syllable or otherwise shorter than the genitive, *zϕ-* occurs to mark the preceding form as genitive but is phonologically bound to the following head as a proclitic. ... When a head noun is longer than a preceding genitive or when the genitive follows the head noun, phonologically independent *zei* occurs to mark the form it follows as genitive’ (Frank 1990: 41).

(39) On the position of possessor NPs (and other adnominal modifiers) relative to the head noun in Turkish, see also Kornfilt (1987: 637): ‘In possessive noun phrases, the possessor precedes the head noun; in “regular” noun phrases, modifiers precede the head. ... The genitive-marked possessor can “scramble” in either direction, while the article and numerals cannot. The adjective is not free to move either, as far as spoken language and written prose are concerned. In poetry, however, an adjective can occur to the right of its head.’

(40) On relative clauses, see e.g. Schwartz (1971); Kuno (1974); Downing (1978); Keenan and Comrie (1977, 1979); Comrie (1981b: chapter 7); Mallinson and Blake (1981: chapter 5); Lehmann (1984, 1986); Keenan (1985); cf. also Maxwell (1979, 1982); Fox (1987).

(41) The phrase ‘adjoined relative clause’ was coined by Hale (1976).

(42) See e.g. Ziv and Cole (1974); Mallinson (1986).

(43) Under certain conditions this ambiguity is resolved (Hale 1976: 79).

(44) There is also this preferred variant with the clause-initial anaphoric element *ŋula* ‘that (one)’: *yankiri-li kutja-lpa ŋapa ŋa-ŋu, ŋula ŋa pantu-ŋu ŋatjulu-ŋu* ‘The emu which was drinking water, that one I speared’ or ‘While the emu was drinking water, then I speared it’.

(45) Cf. also e.g. Downing (1978: 399 ff.), Lehmann (1984: 147 f.), Keenan (1985: 163 f.) and Srivastav (1991) on co(r)relatives; Lehmann (1984: 122 ff.) and Givón (1990: 651 f.) on non-embedded relative clauses. On relative clauses in Hittite, see also Held (1957), Berman (1972), and Justus (1976).

(46) Tamil also has ‘a sort of correlative construction’, which is, however, only rarely used (Asher 1982: 25; see also Hock 1989). I will ignore Abkhaz here, in which a relative clause only follows the noun if the latter is modified by several relative clauses, ‘but such constructions would most naturally be avoided’ (Hewitt 1979: 59).

Apart from the prenominal participial construction, Turkish also has a postnominal relative clause with a finite verb and introduced by *ki* (a conjunction borrowed from Persian). Lewis (1967: 260) writes, however, that ‘the English-speaker composing in Turkish must resist the temptation to translate his relative clauses with the help of *ki*; this use is regarded as alien and is increasingly rare in modern Turkish.’

(47) Notice that this information is taken from Gragg (1972), who writes about Oromo (or Galla) in general, whereas this example is taken from Stroomer’s study on the Boraana dialect (Stroomer 1995). The linker clitic in the example also occurs with the ‘separating element’ -(*t*)*ti*; on the function of this suffix see Stroomer (1995: 113f.).

(48) On double case marking, see Plank (1990, 1995).

(49) Sometimes localizing satellites are regarded as reduced relative clauses (see e.g. Quirk et al. 1985: 1274; Givón 1990: 472).

(50) It has been argued that certain (locative) prepositions in English are to be regarded as predicates (Mackenzie 1992a,b). In this view ‘under’ in ‘from under the house’ is the head of the construction, taking ‘the house’ as its argument.

7 The Underlying Structure of Noun Phrases

7.1. Introduction

This chapter presents a new underlying structure for noun phrases. I will argue that the noun phrase and the clause can essentially be analyzed in a similar fashion and that at some level of abstraction they can be said to share the same grammatical and lexical modifier categories.¹ My analysis has been inspired by Aristotle, who wrote that movement or change ‘pertain[s] exclusively to quality, quantity, and location, each of which embraces contrasts’ (*Physics* 5.2).² Although Aristotle meant to describe physical rather than linguistic phenomena, the notions of Quality, Quantity, and Location appear to be extremely useful in the analysis of linguistic structures as well. In sections 7.2–3 I will first discuss morpho-syntactic manifestations of these three notions at the level of the clause, and, as in the case of the noun phrase, I will take into account both lexical and non-lexical (i.e. grammatical) forms of expressions of the notions Quality, Quantity, and Location. Then I will only briefly return to instances of qualifying, quantifying, and localizing modification in the domain of the noun phrase (section 7.4), since they have already been discussed in detail in Chapters 3–5. Section 7.5 is concerned with symmetrical relations in the underlying structure of NPs and sentences. Section 7.6 gives a formal representation of the underlying structure of the NP and shows how it can be integrated in a general theory of grammar, in particular Simon Dik’s theory of Functional Grammar (Dik 1997).

In section 7.7 I will discuss further parallelisms between NPs and clauses and argue that both the NP and the clause have a descriptive and referential function and that this must be accounted for in the formal representation of the underlying structure. More specifically I will show that in addition to the modifiers in the NP and in the clause that can be divided over the descriptive domains which correspond to three of Aristotle’s basic categories (the *ποιόν* ‘how it is’ = Quality, the *ποδόν* ‘how much/many it is’ = Quantity, and the *ποῦ* ‘where it is’ = Location), there are some modifiers which are essentially concerned with ‘that it is’ – that (p.214) is to say, with the simple fact that referents of NPs and clauses exist (or are not (yet) deemed to exist) in a world of discourse. Let me finally make clear that I will mostly ignore those elements of the clause that have to do with mood/modality and illocution, so that I am only concerned with that part of the underlying clause structure that is called the *predication* in Dik’s theory of Functional Grammar.

7.2. Quality, Quantity, and Location in the clause

It is not difficult to find examples of expressions of *Quality*, *Quantity*, and *Location* in the clause. Tense is of course the ‘grammaticalised expression of location in time’ (Comrie 1985: 9). Tense is a deictic grammatical category in that it relates an event to a reference point (symbolized by • in Fig. 7.1).³ Compare the following sentences:

- (1) Our cat **is eating** a big mouse.
- (2) Our cat **was eating** a big mouse.

The event in (1) is situated in the present because it coincides with the moment of speaking, whereas the event in (2) took place before the moment of speaking and must

therefore be situated in the past.⁴ Since probably all languages have adverbs or adverbial constructions to specify when and where an event occurs it is also easy to find *lexical* expressions of the notion *Location*:

- (3) **This morning** our cat was eating a big mouse **in the garden of our neighbors.**

In (3) the location of the event (*the eating of the mouse by our cat*) is specified more precisely in time (*this morning*) and in space (*in the garden of our neighbors*).

Besides time and space adverb(ial)s, languages also have lexical means to specify how often an event occurs, i.e. lexical expressions of the notion *Quantity*, as in:

- (4) a. Every once in a while our cat catches a mouse.
b. Sometimes our cat catches a mouse.
c. Very often our cat catches a mouse.

Some languages employ grammatical means to indicate how often an event takes place. When reference is made to a single occurrence this is called *semelfactive aspect* and in the case of multiple occurrences this is called *iterative, repetitive*, or (p.215)

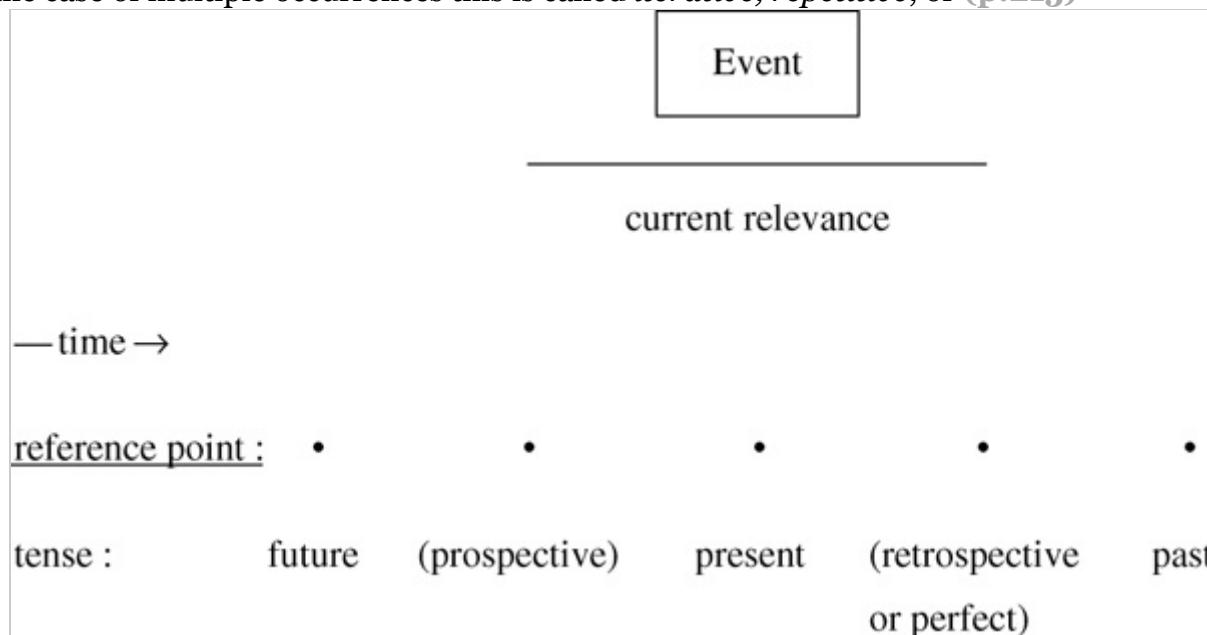


Fig. 7.1. Tense

Hidatsa (Matthews 1965: 158)

- | | | | | | |
|---------------------------------|-------|-----|--------|------|-----------|
| (5) | Wí | i | hírawe | ksa | c |
| | woman | she | sleep | INGR | ITER.MOOD |
| ‘The woman kept falling asleep’ | | | | | |

This sentence also contains a verbal aspect marker (INGR = ingressive aspect) by which the speaker specifies the way in which the verbal property ‘sleep(ing)’ is represented in the temporal dimension in terms of the aspectual features Beginning and Ending. We

saw in sections 2.6 and 4.2.1.1 that verbs designate properties or relations which in principle can all be characterized by these two temporal features (the table is repeated again for the sake of convenience as Table 7.1). When these distinctions are somehow coded by inflectional morphology they are regarded as qualifying (rather than quantifying or localizing) elements because they relate to essential, typical, or characteristic properties. There are of course also *lexical* verb modifiers that express qualitative distinctions, such as adverbs of manner or speed. In the next two examples the adverbs *beautifully* and *slowly* give a further specification of the manner of dancing and the speed of walking:

(6) They danced **beautifully**

(7) The child walked **slowly**

So far we have seen that clauses contain both grammatical and lexical expressions of the three notions that Aristotle deemed relevant for the description of movement/change in physics: quality, quantity, and location. Before I continue with a presentation of these notions in connection with the noun phrase it is useful (p.216)

Table 7.1. Basic aspectual categories of the verb

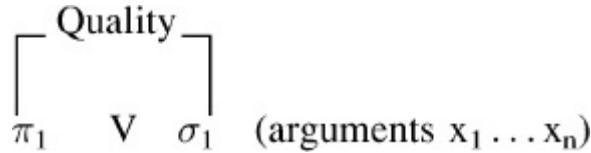
Time	- Beginning	+ Beginning
-Ending	imperfective	ingressive
+Ending	egressive	perfective

to establish that morpho-syntactic manifestations of Quality, Quantity, and Location have different semantic scopes, which in the theory of Functional Grammar (henceforth FG; Dik 1997) is accounted for in the layered organization of the underlying structure of the clause (but note that the model presented here differs from those presented in Hengeveld 1990a and Dik 1997).⁵

7.3. The layered structure of the clause

The FG representation of a linguistic expression is first of all an attempt to account for semantic relationships that exist within the structure of the clause, and the so-called layered model is one of the most important tools to achieve this goal. To put it briefly, in the layered model elements of the linguistic expression that belong together semantically also occur together in the underlying structure of that linguistic expression. For this reason the verbal aspects (symbolized by π_1 = *qualifying clause operator*) and adverbs of manner and speed (and other kinds of *qualifying clause satellites*, symbolized by σ_1) appear closest to the semantic nucleus of the clause: the main predicate V (=verb; I will restrict myself to verbal predicates here). By convention, the formal representation of the underlying structure is written from left to right; operators are placed before and satellites after the nucleus (thus: $\pi_1 V \sigma_1$). Although it is true that semantic relations and structures are often reflected in morphology and syntax (see section 8.5 on *iconicity*), this kind of notation must not be confused with some underlying ('deep') order: the FG representation is a semantic, not a syntactic structure. Since the qualifying operator (π_1) and the qualifying satellite (σ_1) occur in the immediate proximity of the main predicate (V), one could say that they are in the first layer around the nucleus: the *quality layer*.

(8)



(p.217) The next layer accommodates *quantifying clause operators* (such as iterative and semelfactive ‘aspect’, symbolized by π_{2a}) and *quantifying clause satellites* (symbolized by σ_{2a} , i.e. adverbs and adverbials like *sometimes*, *always*, and *every week*).⁶ The scope of *quantifying modifiers* is wider than the scope of *qualifying modifiers*: they influence all the elements on and inside the quality layer. This can be illustrated with the same example from Hidatsa used earlier:

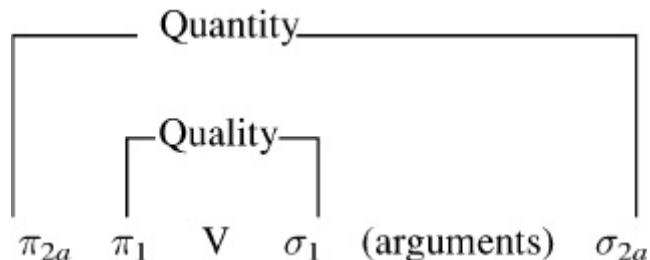
Hidatsa (Matthews 1965)

(9) Wí	i	hírawe	ksa	c
woman	she	sleep	INGR	ITER.MOOD
'The woman fell asleep again and again'				

What is repeated is the falling asleep of the woman, indicating that iterative aspect relates to (i) the predicate (*hírawe* ‘sleep’) with its argument (*wí/i* ‘woman’/‘she’) and (ii) the ingressive marker *ki* (INGR sleep = ‘to fall asleep’).

When we add the *quantifying clause operator* (π_{2a}) and the *quantifying clause satellite* (σ_{2a}) to the structure in (8), we get

(10)



Localizing clause operators (tense markers; symbolized by π_{2b}) and *localizing clause satellites* (time and place adverbials; symbolized by π_{2b}) have the widest scope: they influence everything on and inside the quantity layer. Consider the following sentence:

(11) *The woman fell asleep twice*

The two occurrences of falling asleep by the woman have the same location relative to the deictic center: before the moment of speaking (i.e. in the past). The fact that *Quantity* is in the scope of *Location* and not vice versa is also demonstrated in (12), in which the localizing satellite *this morning* further specifies the temporal location of the two occasions on which the woman fell asleep:

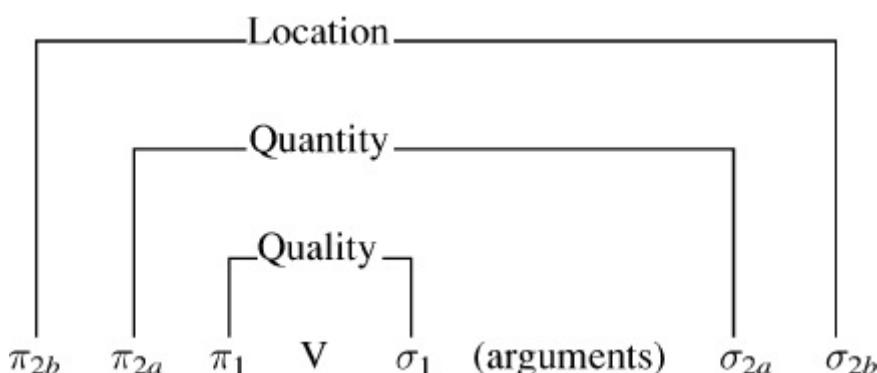
(12) *The woman fell asleep twice this morning*

The repetition has to take place within the time interval as specified by the localizing satellite *this morning*; therefore it is not possible to say:

- (13) *At a particular moment the woman fell asleep twice

(p.218) The time span of a moment is too short to fall asleep twice. And if *at a particular moment* would be in the scope of *twice*, the woman should have fallen asleep at two particular moments and this is clearly not contained in the meaning of (13). Adding *localizing clause operators* (π_{2b}) and *localizing clause satellites* (σ_{2b}) to the structure in (10), we get the following (simplified version) of the layered model of the clause (or rather, as mentioned earlier, that part of the clause that is called ‘predication’ in FG);

- (14)



Since this is a general structure that is designed to accommodate linguistic expressions from any natural language, certain slots may remain unused in a particular language. For example, most languages do not employ *quantifying clause operators* (inflectional markers of semelfactive and iterative aspect) but only express quantifying notions through lexical means, i.e. quantifying clause satellites such as ‘always’, ‘every week’, and ‘twice a day’.

In the next section I will argue that it is possible to construe a similar structure with hierarchically nested layers for noun phrases.

7.4. The layered structure of the noun phrase

In Chapters 3–6 I have demonstrated that we also find morpho-syntactic manifestations of the three Aristotelian notions *Quality*, *Quantity*, and *Location* in the noun phrase. If a language has adjectives, they are most commonly used as *qualifying satellites* (section 4.3) in that they typically specify a characteristic, more or less inherent property (or *quality*) of a referent, such as its size, value, weight, age, or color –as in:⁷

- (15) this **big/expensive/heavy/old/brown** table

In section 4.2 I argued that nominal aspect markers (or *qualifying operators*) are grammatical expressions of the notion *Quality* in the NP. Since they pertain to the type of entity that is defined by the noun in that they specify in more detail how the nominal property is represented in space (in terms of the features Shape and (p.219) Homogeneity), they too relate to an inherent characteristic of the referent of the noun phrase. Thus, in (16a) the noun is ambiguous between defining a singleton

and a collective set, whereas in (16b) the nominal aspect marker indicates that the referent involves a collective entity.

Oromo (Stroomer 1987: 77)

- | | |
|---------------|---------------------------|
| (16) a. saree | 'dog/dogs' (unmarked set) |
| b. sareellee | 'dogs' (collective set) |

Quantifying operators and *satellites* in the NP (i.e. grammatical and lexical expressions of number and cardinality) were discussed in Chapter 5. The English example below contains a grammatical expression of cardinality, whereas in the other example from Boumaa Fijian the numeral is categorized as a verbal element (note that *gone* is a transnumeral noun):

- (17) *the two children*

Boumaa Fijian (Dixon 1988: 144)

- | | | | |
|---|----------|-----|------------|
| (18) e | tolu | a | gone |
| 3SG.S | be_three | ART | child(ren) |
| 'three children' [lit. the children they are three] | | | |

Finally, in Chapter 6 I argued that several types of noun modifiers have in common that they specify the location of the referent of the noun phrase, thus making it possible for the addressee to correctly identify the intended referent. Adnominal demonstrative pronouns are of course grammatical manifestations of the notion *location (localizing operator)*, whereas (restrictive, identifying) relative clauses, possessor phrases, and other kinds of adpositional or case marked NPs are good examples of lexical manifestations of the notion *location (localizing satellites)*.

- (19) **this book**

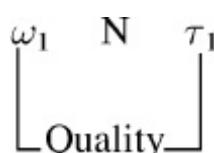
- (20) **the book on the table** [location in space]

- (21) **the book I gave you (yesterday)** [location in time]

- (22) **John's book.**⁸

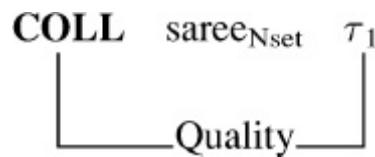
As announced in the beginning of this section, we find the same hierarchical scope relations in the noun phrase as in the clause. Qualifying operators ω_1 (nominal aspect markers) and qualifying satellites τ_1 (typically adjectives, if a language has them) are closest to the head of the noun phrase because they only relate to the entity as defined by the noun. (p.220)

- (23)



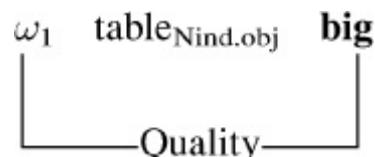
Recall that each of the two nominal aspect markers specifies what kind of set is defined by the noun (a singleton set, containing a singular object, or a collective set) and that adjectives (or other formal manifestations of quality satellites) typically specify more or less *inherent* or characteristic properties that come with the kind of entity as defined by the noun. Since in this study we are only considering nouns used in connection with first order (spatial) entities, such properties may involve such domains as size (big, small), color (black, blue), weight (light, heavy), etc. Thus, in this simplified mode of representation Oromo *saree-lle* [dog_{Nset}-COLL] ‘dogs’ would be represented as follows (recall, once again, that this is an incomplete and strongly simplified version of the underlying structure; the complete representation, which serves as the input for morphological and syntactic rules in the expression component of the FG model, is given in section 7.6 below):

(24)



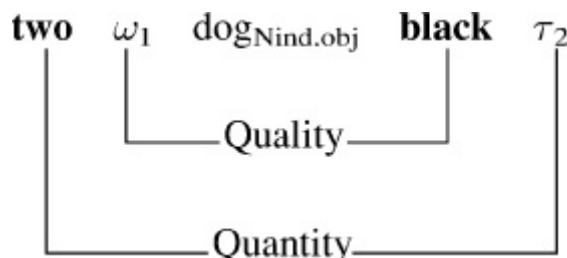
And ‘(the/a) **big** table’ as

(25)



By contrast, Quantity (number, cardinality) and Location are notions that relate to non-inherent or external features of a referent and have in their scope the entity with all its characteristics as specified by the noun and its qualifying modifiers ω₁ and τ₁ (if present). For example, in ‘**two** black dogs’ it cannot be the case that only one of the two dogs is black.

(26)



Similarly, in ‘**these** black dogs’ the speaker only refers to the dogs that are black at the place specified by the localizing modifier *these* (proximate from the speaker’s perspective).⁹ (p.221)

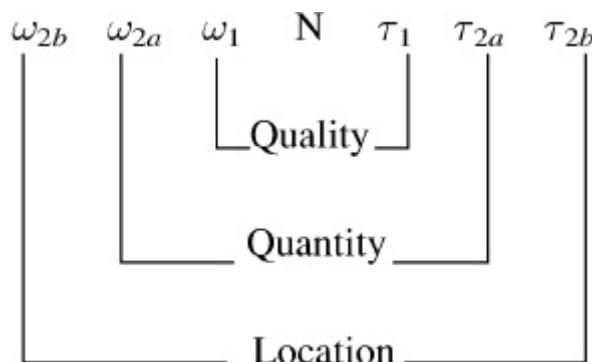
(27)



This shows that qualifying satellites (τ_1) are in the scope of quantifying and localizing modifiers. Further evidence is provided by the influence of numerals (quantifying operators ω_2) on the nominal aspect marker (qualifying operator ω_1).

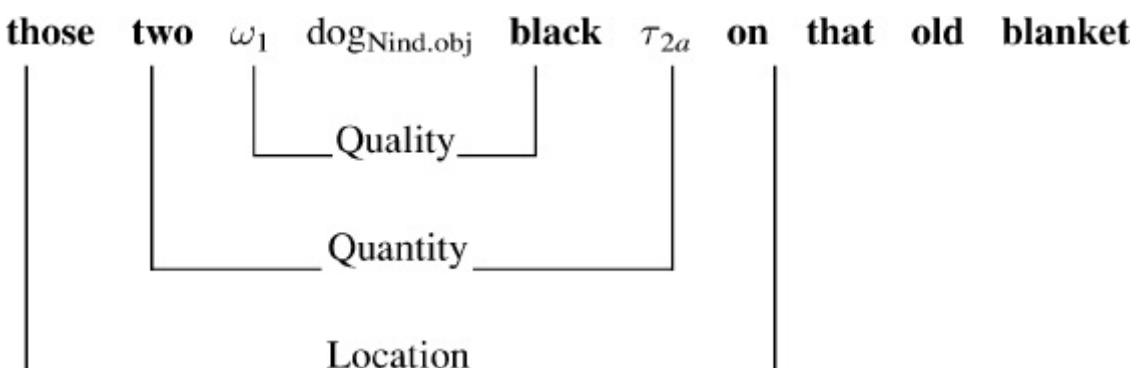
Regarding the differences in scope between quantifying and localizing modifiers, we find that, as in the clausal domain, quantifying modifiers are inside the scope of Location modifiers. The latter specify the location of the entity with all its quantitative and qualitative properties. In other words, the number or cardinality of the referent of an NP is entirely contained within the region indicated by the localizing modifiers ω_{2b} and/or τ_{2b} .

(28)



Thus, in ‘those two black dogs on that old blanket’ the old blanket in the area at some distance from the speaker is the place of both black dogs.

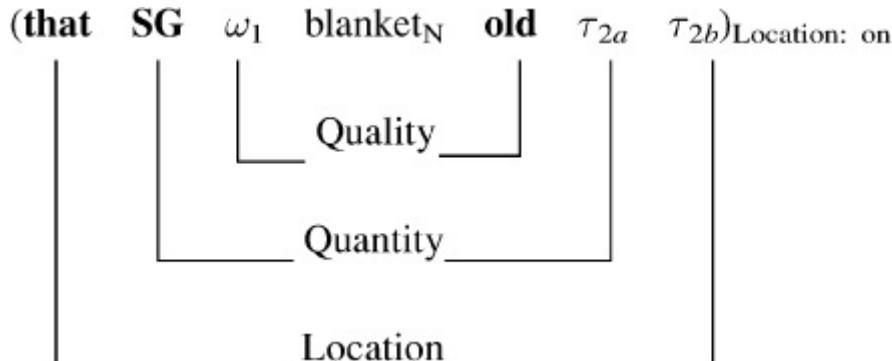
(29)



If it were the other way around, so that Location were in the scope of Quantity, the quantifying modifiers would not only specify the number or cardinality of the referent as defined by the noun *cum* quality modifiers (black dogs), but also the number or cardinality of the location (those/that old blanket) and this is clearly not the case: there are two dogs but there is only one blanket, which serves as the single location of both dogs.

So what we see is that the quantifying modifier in the matrix NP (*two*) does not extend its influence to the localizing operator *those* and the localizing satellite (p.222) (*on*) *that old blanket*, as a matter of fact, this modifier has its own nested layers with slots for qualifying, quantifying, and localizing modifiers:

(30)



Since the underlying structure allows for recursion (in principle, any satellite τ has its own modifier slots), it is theoretically possible to have a long string of satellites (phrasal modifiers). This possibility is usually only exploited for identification purposes (involving the τ_{2b} slot of underlying structure), as when the speaker assumes that the referent cannot be identified without providing the hearer with some extra clues. For example, the modifier ‘on that old blanket’ can itself be modified by another location satellite (such as ‘in front of the fireplace’):

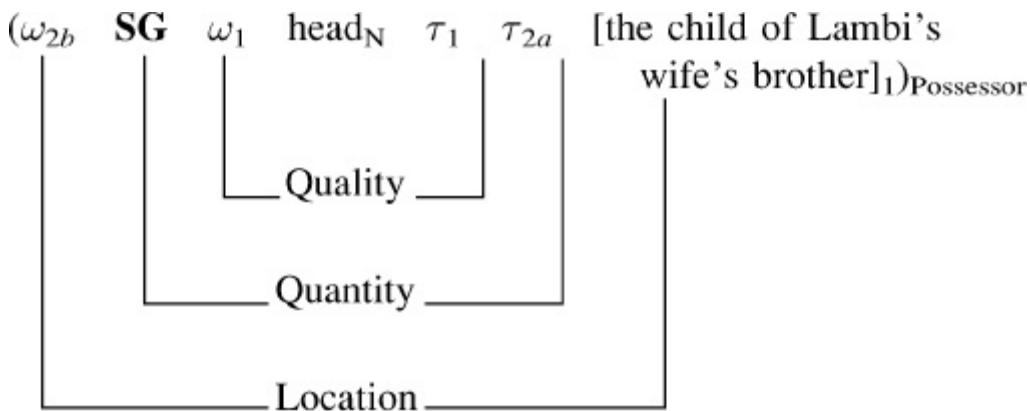
(31) Those two black dogs [on that old blanket [in front of the fireplace]Location]Location

An example from Babungo of such a complex NP was given earlier in Chapter 6 (as well as in Chapter 1). Here *tó* ‘head’ (‘hair’) is identified by reference to ‘Lambi’ (the so-called Identifying Entity or IE; see Chapter 6) via three other entities: ‘Lambi’s wife’, ‘Lambi’s wife’s brother’, and finally ‘the child of Lambi’s wife’s brother’. In the representation below square brackets identify the embedded localizing satellites (the place of (in)definiteness in the underlying NP structure is discussed below):

Babungo (Schaub 1985: 76)

(32)	(ŋwé bá)	tó	[wée	[wéenshú	[zú [Làmbí]4]3]2]1
	(she weave:PRES)	head	child	brother	wife Lambi
‘(she is plaiting) the hair of the child of Lambi’s wife’s brother’					

(33)



- (34) $(\omega_{2b} \omega_1 \mathbf{SG} \text{child}_N \tau_1 \tau_{2a} [\text{Lambi's wife's brother}])_{\text{Possessor}}$

(p.223)

- (35) $(\omega_{2b} \mathbf{SG} \omega_1 \text{brother}_N \tau_1 \tau_{2a} [\text{Lambi's wife}]_3)_{\text{Possessor}}$

- (36) $(\omega_{2b} \mathbf{SG} \omega_1 \text{wife}_N \tau_1 \tau_{2a} [\text{Lambi}]_4)_{\text{Possessor}}$

To sum up, in an NP like ‘those two black dogs on the carpet’ it is only the dogs that are black, not the quantity or the location. And the quantifying modifier *two* specifies the number of black dog entities, not the number of locations; and both *those* and (*on*) *that old blanket* specify the location of dog entities with all their qualitative and quantitative properties as given in the NP. From this we may conclude that the underlying structure of the NP and that of the clause (or, better, the *predication*) are each other’s mirror image (see Fig. 7.2).¹⁰ Recall that, particularly in the case of lexical modifiers (clause and NP satellites σ and τ) there is no one-to-one relation between form and function. Especially relative clauses and adverb(ial)s are very versatile in that they are employed as qualifying, quantifying, and localizing modifiers. For example, although the qualifying noun modifier (τ_1) typically takes the form of an adjective (if a language has adjectives), it may also be realized by a (non-restrictive, non-localizing/non-identifying) relative clause. Compare

- (37) He gave her a [rather cheap] τ_1 ring

- (38) He gave her a ring [that was actually rather cheap] τ_1

On the other hand, adjectives can also be used for identifying purposes, although the fact that they have (contrastive) focus may indicate that this is not their normal function (Chapter 6).

- (39) No, I would like the RED τ_{2b} pepper (not the green one)

7.5. Symmetry between NPs and clauses

How can we explain parallels in the underlying structure of NPs and clauses? Possible answers to this question seem to revolve around the way that humans conceive entities. Several recent studies have argued that our conceptual system is largely metaphorical in nature.¹¹ In this view new or complex entities are conceived in the same way as familiar, more basic entities, and since there are many examples which show that spatial metaphors are used to express temporal and other non-spatial notions, it is generally assumed that spatial conception plays a fundamental role in human cognition (see section 6.1 on the *localist hypothesis*). Witness (Levinson 1992: 6):

(40) *before* the altar [space] vs. *before* tomorrow [time]

(41) *distant relative; closest kin* (kinship)

(p.224)

Grammatical expression of Quality (π_1),
Quantity (π_{2a}), and Location (π_{2b})
in the clause

Lexical expression of Quality (σ_1),
Quantity (σ_{2a}), and Location (σ_{2b})
in the clause

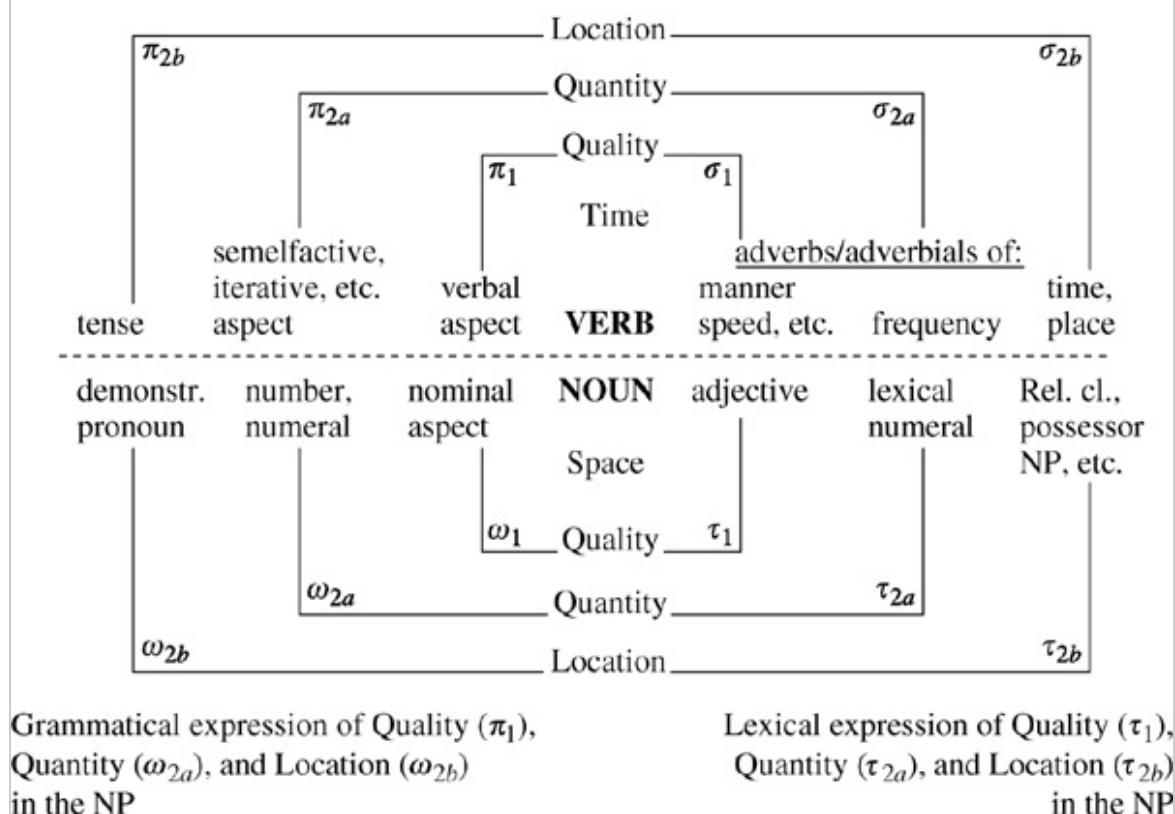


Fig. 7.2. Symmetry in the underlying structure of the clause and the NP

(42) *high* notes; *flat* notes (music)

(43) *high* and *low* numbers (mathematics)

(44) to feel *depressed* or *alone* (emotions)

(45) the *ascent* of man; the *peak* of his career; his *circle* of friends

Thus one could hypothesize that the observed parallels between the underlying structure of NP and clause might be due to the fact that temporal entities are understood in terms of (cognitively less complex) spatial entities.¹² In other words, the way that humans construe referents of clauses (events) is modeled after the way that they construe referents of NPs, because the latter (spatial entities) are conceived more clearly.

(p.225) 7.6. Formal representation in Functional Grammar

Since the standard representation of NPs in Functional Grammar will constitute the point of departure for a more formal version of the underlying structure proposed in this book, a brief review of the main aspects of the way NPs are represented in Functional Grammar (FG) is in order before we proceed. In Simon Dik's theory of FG, NPs are represented according to the following general schema (Dik 1997: 132):¹³

$$(46) (\omega x_i: \varphi_1(x_i): \varphi_2(x_i): \dots: \varphi_n(x_i)) [n \geq 1]$$

As in the representation presented earlier in this chapter, ω stands for various kinds of *NP operators* (or *term operators*), i.e. grammatical categories such as (*in*)definiteness (i or d) and *number* (1 'singular' or m 'plural') or *cardinality* (any cardinal number); x_i is the term (NP) variable, 'ranging over the set of potential referents of the term', and $\varphi_1 \dots \varphi_n$ signify (nominal, adjectival, or verbal) predicates, also called *restrictors*, because they successively restrict the range of possible referents of the NP.

Each φ is an open predication in x_i , predicating a property of referent x_i , or a relation between x_i and maximally two other referents (e.g. x_j and x_k). The first restrictor (φ_1) is normally the head noun; other restrictors can take a variety of forms (e.g. adjective, possessor phrase, relative clause). The colon between two restrictors indicates that the information to the right gives a specification of, or a restriction on, the possible values of x_i as it has been specified at that point. In previous chapters non-first restrictors were referred to as *satellites* (more on this below).

To give an example, the underlying structure of the simple NP 'the two black dogs' is as follows:

$$(47) (d2x_i: \text{dog}_N(x_i): \text{black}_A(x_i))$$

This is to be paraphrased as 'definite two entities x_i such that the nominal property "dog" applies to x_i , such that the adjectival property "black" applies to x_i '.

A complex NP such as 'the two black dogs on that blanket', which contains the *embedded* satellite 'on that blanket', would look like this (rem = remote):

$$(48) (d2x_i: \text{dog}_N(x_i): \text{black}_A(x_i): \{(d_rem1x_j: \text{blanket}_N(x_j))_{\text{Loc}}\}(x_i))$$

Here the satellite 'on that blanket', which is actually a *term predicate* with the semantic function Location (Loc), serves as the third restrictor of the referent symbolized by x_i .

(p.226) The next example has as the second restrictor a verb heading a relative clause, which refers to an event (symbolized by the e-variable; R = relative operator, Ag = Agent, Go = Goal).

$$(49) (\text{dm}x_i: \text{dog}_N(x_i): [\text{Past } e_i \text{ attack}_v (Rx_i)_{\text{Ag}} (\text{dm}x_j: \text{child}_N(x_j))_{\text{Go}}]) \text{ 'the dogs that attacked the children'}$$

In the complete FG representation of a linguistic expression each NP would be provided with a specification of its semantic and possibly its pragmatic and syntactic role.¹⁴ A complete representation of an utterance would also include layers that provide slots for propositional and illocutionary operators and satellites. When the underlying structure

contains all those elements and relations that are needed to arrive at a correct semantic and pragmatic interpretation, it is the input for the expression rule component. Expression rules are responsible for the form of the constituents (inflectional morphology), the linear order in which they will be expressed (syntax), and the prosodic contour of the expression (phonology); for a general overview of the theory of Functional Grammar I refer to Dik (1997).

As has already been indicated above, in the model of the noun phrase that is proposed here the various NP constituents (restrictors, operators) are associated with one of three hierarchically nested *layers* of the underlying NP structures (Fig. 7.3).¹⁵

The structure in (50) is the complete formal representation of this NP-model in the FG format. Besides the variables x , ω , and τ , this schema also contains the

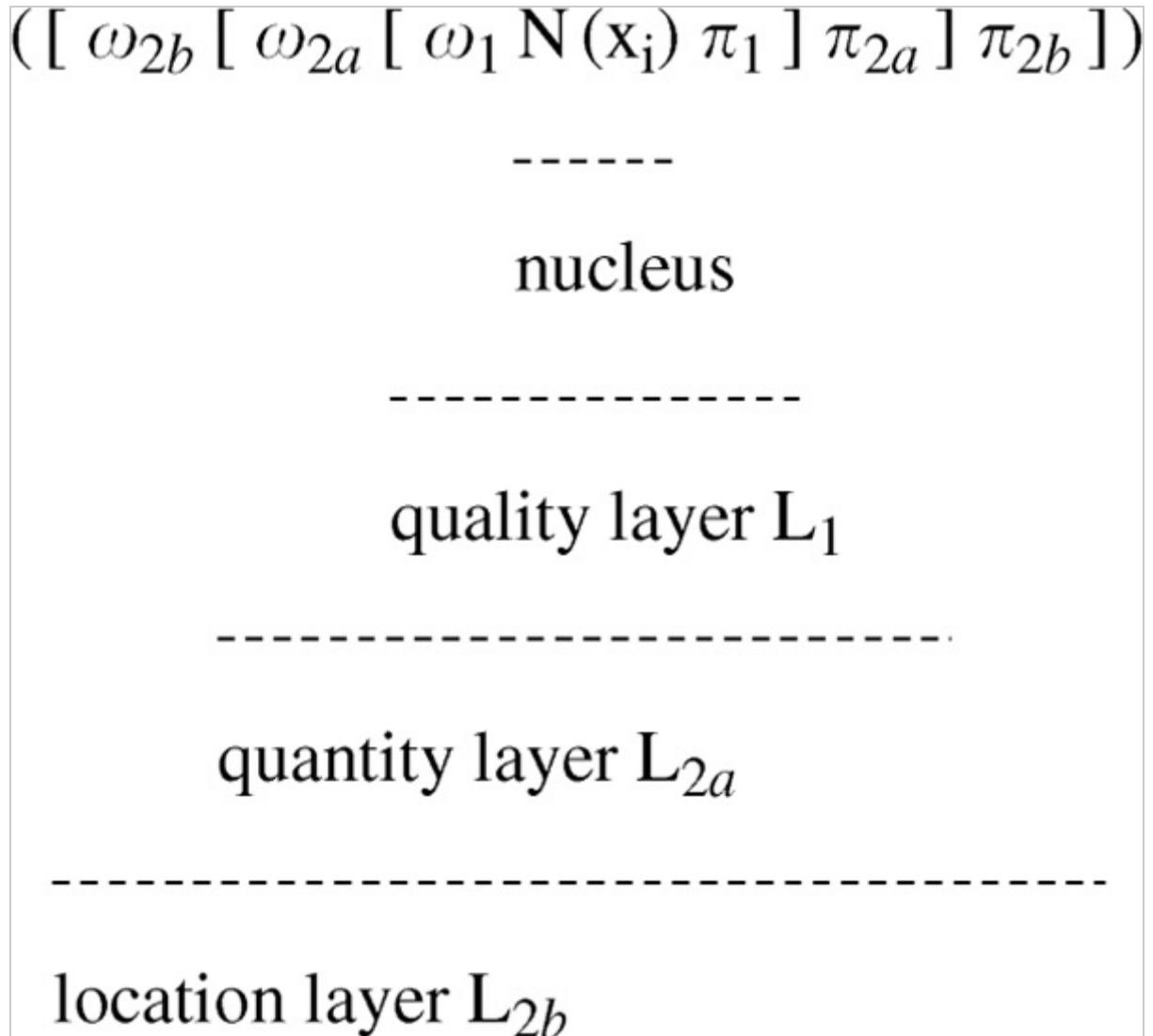


Fig. 7.3. The hierarchical structure of the NP
 (p.227) referent variable X , the discourse operator Ω the discourse satellite T , and the predicate variable f , all of which are discussed below.¹⁶

(50) $(\Omega[\{\omega_{2b}[\omega_{2a}[\omega_1[\{N(f_i)\}(x_i)]_{L0}\{\tau_1(f_j)\}]_{L1}\{\tau_{2a}(f_k)\}(x_i)]_{L2b}\{\tau_{2b}(f_1)\}(x_i)]_{L2b}(f_m)\}(X_i)]\{T_i(f_n)\}(X_i))$

Or, in a somewhat more transparent version (without the predicate variables):

(51) $(\Omega [\omega_{2b} [\omega_{2a} [\omega_1 [N(x_i)]_{L0} \tau_1]_{L1} \tau_{2a}(X_i)]_{L2a} \tau_{2b}(X_i)] T_i(X_i))$

Ω = discourse operator (e.g. *(in)definiteness* or *(non-)specificness*; this operator has the widest scope and is concerned with discourse properties of the referent)

X = referent variable (symbolizes the referent of the NP)

T = discourse satellite (lexical modifier concerned with discourse properties of the referent; the T satellite takes the referent, symbolized by X_i , as its argument)

ω = NP operator; ω_1 has the nominal predicate in its scope; ω_{2a} has in its scope the quality layer L_{1a} ; ω_{2b} has in its scope the quantity layer L_{2a}

N = (head) noun

f = predicate variable

x = NP or term variable (symbolizes the noun phrase that the speaker uses to refer to the discourse entity, i.e. the speaker's description of the referent of the NP)

$\tau(f)$ = NP or term satellite; the argument of qualifying term satellite $\tau_1(f_j)$ is the property designated by the noun (f_i); the argument of quantifying term satellite $\tau_{2a}(f_k)$ is the spatial entity x_i as defined by material in the quality layer L_1 ; the argument of localizing term satellite $\tau_{2b}(f_l)$ is the spatial entity x_i as defined by material in the quantity layer L_{2a} .

The relevance of the predicate variable 'f' was anticipated by Dik (1989: 50), the rationale being that, strictly speaking, every lexical constituent is a specification of a *predicate* or *content variable*. Since then it has been used in the representations of various kinds of structures.¹⁷ For example, in *Peter was happy and so was Mary*, 'so' is an anaphorical reference to the predicate '(be) happy', as shown in the simplified representation of this sentence below (A stands for 'anaphoric'; for more details see Vet 1986; Keizer 1992; Dik 1997):

(52) ([Past e_i : $happy_{Adj}(f_i)$ (Peter)] & [Past e_i : A(f_i) (Mary)])

In the alternative NP model the f-variable of the head noun (f_i) serves as the argument of a qualifying satellite, τ_1 (i.e. $\tau_1(f_j)(f_i)$); note that the qualifying satellite τ_1 has its own predicate variable: (f_j)), because these satellites relate (p.228) directly to the property designated by the nominal predicate (hence the alternative name *predicate satellite*)¹⁸

Apart from the layered organization, there is another important difference between Dik's NP model and the underlying NP structure proposed here. This concerns the use of the referent variable X, which has its own operator and satellite slots (Ω and T, respectively). The main reason to distinguish between the linguistic construction (i.e. the NP) and the referent of that linguistic construction is that it clearly brings out the dual function of the NP. On the one hand NPs give a 'physical' description of a referent regarding its qualitative ('how'), quantitative ('how many'), and positional ('where') properties. This descriptive aspect of the NP, which is concerned with properties of the referent as a spatial entity, is symbolized by the x-variable in the model presented here. On the other hand, NPs are referring expressions and the referent of an NP is an entity that is talked about and figures in the world of discourse. This constitutes the referential side of the NP and is symbolized by the X-variable in the NP model proposed

above. Thus, the x -variable has received a different interpretation in the alternative model of the NP that I propose here. The x -variable does *not* stand for the intended referent, as in Dik's model of the NP; referents are now symbolized by the X -variable.

Perhaps one could say that, essentially, before ' x ' is part of an underlying NP structure, it simply symbolizes 'space', the spatial dimension, which through lexical and grammatical specifications acquires certain properties.¹⁹ Initially a nominal predicate *projects* in the spatial dimension the most typical property of the entity, e.g. ' $\text{book}_{N_sing.\text{obj}}$ ' or ' $\text{chair}_{N_sing.\text{obj}}$ '. Since nouns are lexically coded for a certain *Seinsart* (like ' $N_sing.\text{obj}$ ') what is also specified at this stage is *how* the property projected by the head noun in question is represented in the spatial dimension in terms of the features Shape and Homogeneity (Chapter 2). Then qualifying, quantifying, and localizing operators and satellites (if present) specify further descriptive information of the referent. These modifiers, however, only concern the physical properties of the referent of the NP, and do not inform the addressee about the status of the referent in the world of discourse.²⁰

(p.229) Thus, whereas most modifiers in the noun phrase are *descriptive modifiers*, which can be distributed over semantic domains which correspond to three of Aristotle's basic categories (the $\piοιόν$ 'how it is' (Quality), the $\piοδόν$ 'how much/many it is' (Quantity), and the $\piοῦ$ 'where it is' (Location)), there are also modifiers which relate to the occurrence of the referent in the world of discourse (*discourse modifiers*) and which are discussed in more detail in section 7.6.1 below.

Another reason to distinguish between the linguistic expression proper and the referent of that expression is that we can now clearly account for the fact that different NPs (x_i and x_j) can refer to the same referent (X_i), as in this well-known example (Frege 1892):

(53) *The Morning Star* (X_i, x_i) *is the Evening Star* (X_i, x_j).

The practical need for this distinction is demonstrated by the fact that it is now possible to specify in the formal representation whether anaphoric reference is made to an NP (x) or to the referent of an NP (X); compare (A = anaphoric operator):

(54) A: My neighbor (X_i, x_i) just saw a black cat and now the superstitious old fool (X_i, x_j) believes he (A X_i) is in for some bad luck today.

B: Why do you call him (A X_i) that (A x_j)?

In this example *he/him* refers to the discourse entity (X_i), whereas *that* refers to one of the speaker's descriptions of that entity, the NP indexed as (x_j).

7.6.1. Discourse modifiers

In the new NP model presented above the referent is symbolized by the referent variable X , which has its own operators and satellites. These modifiers are concerned with the status of the referent as a discourse entity. It was announced in Chapter 6 (note 2) that articles, which are generally regarded as markers of 'weak deixis', are better analyzed as *discourse operators*. In a given world of discourse, referents are either

identifiable or not identifiable and articles typically indicate whether the speaker regards the referent of an NP as identifiable (definite) or not identifiable for the hearer (indefinite). There is basically only one reason why a referent is not identifiable: because it has not been *properly* introduced in the shared world of discourse of the Speaker and the Addressee(s), as in:

(55) *Why don't you go out and buy a newspaper?*

(p.230) I used the manner adverb 'properly' because sometimes the referent of an indefinite NP has already been introduced in some vague fashion, for example, as a member of a larger group (as in 'I bought some books'). Although we can subsequently refer to the complete same set of entities with a definite NP (*The/these books were ...*), we have to use an indefinite NP to refer to any subset. This is also exemplified in (56), where the referent in question (*a book*) is presented as unidentifiable, despite the fact that it has been singled out from a larger number of entities of the same kind that was introduced earlier (*some old books*).

(56) *On the table were some old books and magazines. I picked up a book and started reading.*

There are, however, many reasons why a referent of a (first order) NP is presented as identifiable by the Speaker; for example, because it was introduced earlier in the discourse (previous mentioning), because the ontological correlate of the referent is available in the physical context, or because in the given context the referent is believed to be a unique entity.

Referent has been mentioned earlier:

(57) *I just bought a book and an agenda. Surprisingly, the book was much cheaper than the agenda.*

Referent is available in physical context:

(58) *Now tell me—what do you see on the monitor?*

Referent is unique in given context (Amsterdam):

(59) *We decided to meet in front of the city hall.*

We saw in Chapter 6 that definite NPs can also introduce new discourse referents through the back door, as it were, as when their occurrence in the world of discourse is presupposed. In such cases the existence of a referent can be inferred through its Identifying Entity (here 'the Pentagon'; see also the introduction to Chapter 6 and the Babungo example above).

Referent's existence is presupposed (apparently there is a book shop opposite the Pentagon):

(60) *The book shop opposite the Pentagon has an excellent selection of national and international newspapers.*

Discourse satellites are *lexical modifiers* that provide the addressee with information about the referent as a discourse entity. They typically (further) specify when or where a referent was introduced in the previous (spoken or written) discourse, such as the Dutch modifier *zojuist genoemde* ('recently mentioned'):

(61) De zojuist genoemde personen krijgen allemaal een prijs
the just mentioned persons get all a prize

'The persons whose names have just been mentioned will all get a prize'

(p.231) This modifier is different from the ones discussed in Chapters 4–6 in this book, because it does not tell us anything about the ‘how’ (quality), the ‘how many’ (‘quantity’), or the ‘where’ (location) of the referent. Instead it indicates that the entities involved already exist in the world of discourse and are now collectively being referred to again. One might argue that in some sense *zojuist genoemde* tells us something about the ‘where’, in that it points to a certain place (or moment) in the discourse. However, whereas *localizing satellites* specify the location of the referent in the world of discourse, *discourse satellites* such as *zojuist genoemde* ‘just mentioned’ (but also modifiers such as ‘the former’, ‘the latter’) relate to the location where the referent was mentioned before in the actual conversation (*discourse deixis*).

Modifiers such as ‘(the) same’ and ‘(the) other’ as in

(62) *I gave her the same/the other book*

also belong to the category of discourse modifiers in that they, too, are concerned with discourse properties of the referent. In the case of ‘(the) same’ the addressee is explicitly instructed to locate a certain referent that was mentioned a little earlier in the conversation. In the case of ‘(the) other’ the Addressee should identify the second member of a set that has already been established as a discourse referent. By contrast, the modifiers ‘(a) different’, ‘another’ (as in: *I gave her another book*) are an instruction for the addressee to *construe* a new referent of the same type (*book*) as a discourse referent that already exists as a distinct referent in the discourse world and was mentioned not long before.²¹

7.7. More parallels between NPs and clauses

In the previous section I have argued that lexical NPs have a dual function. On the one hand they are descriptive phrases specifying certain qualitative, quantitative, and/or locative properties of an entity (recall that I am only concerned with NPs referring to spatial objects); on the other hand they are *referring expressions* in that they point to an entity in the discourse world.²² In view of this new distinction in the underlying structure of the NP we can now say similar things about clauses, making it possible to extend the parallels between the underlying structure of NPs and clauses discussed above.

Just like a lexical NP, a clause both gives a description of an entity *and* refers to that entity in the world of discourse. Thus at the level of the clause we can also distinguish between the speaker’s description of an entity (the linguistic structure: the clause) and the referent of the speaker’s description (the event). I will (p.232) use the letter E as the referent variable (symbolizing the event), and use Dik’s e-variable to symbolize the speaker’s description of the event, i.e. the clause (or, rather, that part of the underlying clause structure that is called *predication* in FG). The following dialogues illustrate the usefulness of this distinction for anaphoric purposes.

(63) A: Fred saw [that Harry was talking to the police this morning] (E_i, e_i).

B: Yes, I saw that (A E_i) too.

(64) A: Fred saw [that Harry was talking to the police this morning](E_i, e_i).

B: That(A e_i is not (a good description of) what really happened; the police were actually interrogating him.

In (63) person B uses ‘that’ as an anaphoric reference to what Fred saw, i.e. the event (what one sees is an event), whereas in (64) person B uses ‘that’ to refer to Fred’s description of the event.²³

Like the X-variable, the E-variable has its own operator and satellite slots for grammatical and lexical modifiers that provide information about the referent as a discourse entity. *Discourse operators*, symbolized by Π in the formal representation below, are *actual* (or *realis*) and its negative counterpart *non-actual* (or *irrealis*). They indicate whether or not event E as described in predication e has *actually* occurred or is *actually* occurring.²⁴ When an event is actual it is anchored (or *grounded*; cf. Chapter 6) in the world of discourse in that it can be located in time and the clause will be tensed (provided tense is a grammatical category in the language in question).²⁵ Conversely, if the speaker refers to a non-actual event, it will not be possible to locate the event in time, and the clause cannot be tensed. An event may be non-actual for several reasons; for instance, because the predication referring to this event is expressed as a command (referring to an event to be brought about by the addressee), because the possible occurrence of the event denoted by the predication is hoped, feared, or doubted, or because the speaker wants to indicate that he or she has no direct evidence for the actual occurrence of (p.233) the event.²⁶ In the latter case one of the *evidential moods* may indicate that the speaker has inferred the (possible) occurrence of the event (*inferential mood*) or that someone else told him about the event in question (‘hearsay’; *quotative* or *reportative mood*). For this reason non-actual (non-tensed) predications are commonly under the scope of certain modal or illocutionary operators. In West Greenlandic, for example, tense affixes do not occur in the imperative/optative mood, at least not ‘in their strictly temporal sense’ (Fortescue 1984: 275), and in Kobon ‘the tense distinctions between simple past, remote past, present, and future are made only in the indicative mood’ (Davies 1981: 168).²⁷

Takelma and Palauan (in the sample, for instance, Bukiyp and Korean; cf. section 4.3.3) belong to the languages that appear to employ special morphemes to mark the distinction between actual/non-actual.²⁸ The first example from Takelma has the verb with a distinct realis stem, whereas the verb stem in the second sentence has the inferential suffix -k'- and appears in the form that is used for all non-actual moods:

Takelma (Chung and Timberlake 1985: 244; from Sapir 1922: 158)

(65) Menà yap'a t' omõ-k'wa

bear man kill(REALIS)-3HUMAN_OBJECT

‘The bear killed the man’

(66) Menà yap'a dõm-k'wa-k'

bear man kill(IRREALIS)-3HUMAN_OBJECT-INFERENTIAL

‘It seems that the bear killed the man/The bear must have, evidently has, killed the man’

The other two examples are from Palauan, which ‘uses its single irrealis mood in the non-past tense for commands, as well as exhortative and desiderative senses’.²⁹

(p.234) Palauan (Chung and Timberlake 1985: 248–9; from Josephs 1975)

(67) Ku-rael el mo er a blik

1SG.IRR-travel COMP go LOC ART my_house

'Perhaps I should go home'

(68) Mo-lim a kərum!

2SG.IRR-drink ART your_medicine

'Drink your medicine!'

There are, however, also many languages that seem to do without a distinct (non-)actuality marker. Presumably this is so because all real *tense* markers imply *actuality*, so if a language marks *tense* there is no need to also mark *actuality* (see in this context also Comrie 1985: 51 on Burmese and Dyirbal).³⁰ Conversely, the reason why *non-actuality* often receives no overt coding is due to the fact that this is in many instances implied by a particular mood (e.g. dubitative) or illocution (e.g. imperative).

It is clear that what the *discourse operators* 'actual'/'non-actual' and 'definite'/'indefinite' have in common is that they relate to the occurrence of the referent in the world of discourse. The operators 'Definite' and 'Actual' in the formal representation of the underlying structure signal that the speaker wants to indicate that the referents of these phrases have a place in the world of discourse. Conversely, the operators 'Indefinite' and 'Non-actual' in the underlying structure are an indication for the addressee that the referents of these phrases do not (yet) have an established (identifiable) place in the world of discourse. In the case of a first order NP the hearer should interpret the *indefiniteness operator* as an instruction to construct a new spatial entity as a discourse referent. The *non-actuality operator* in the underlying structure of the clause also instructs the addressee to construe a new entity as a discourse referent (in this case: an event), but in addition the *non-actuality operator* tells the addressee that this event must not be 'grounded' or 'anchored' as a real event in the world of discourse.³¹

So what we see is that 'definite' and 'actual' have a similar function in that they signal that the entities they refer to (already) *exist* in the world of discourse (or that their existence is presupposed). Their negative counterparts 'indefinite' and 'non-actual' on the other hand have in common that the entities they refer to do not exist (or do not exist yet) in the world of discourse as identifiable or actual ('anchored') entities (see Table 7.2).

(p.235)

Table 7.2. Discourse operators Ω and Π

Referent (already) exists in discourse world		Referent does not exist in discourse world (yet)
Noun phrase	$\Omega = \text{definite}$	$\Omega = \text{indefinite}$
Predication	$\Pi = \text{actual} (\text{realis})$	$\Pi = \text{non-actual} (\text{irrealis})$

The notions *specificness* and *genericness*, which are not considered in any detail in this book (see sections 1.5.4 and 3.3.2.4.1), are probably also best categorized as discourse operators Ω , since both categories have to do with discourse properties of a referent.³² In the case of specific reference (+SP), the (actual or quoted) speaker has a certain referent in mind, although the addressee may or may not be able to identify this referent. In the case of non-specific reference (-SP), the speaker does not have a particular referent in mind in that any entity that fits the description will do. Although there are languages, such as Samoan (Mosel and Hovdhaugen 1992: 259) and Port Sandwich (Crowley 1985: 154), that have distinct markers to express specific and non-specific reference (see also Greenberg 1978a on Stage II articles in African languages), most languages have no special marker for NPs that have (non-)specific reference and instead use, for example, a definite or an indefinite article for this purpose. This has the disadvantage that there is often a certain degree of ambiguity, as illustrated in the two following sentences. Compare:

(69) *This morning I saw that Max was talking to a taxi driver* (+SP).

(70) *I want to talk to a taxi driver* (+/-SP).

Whereas in (69) there is no doubt that the speaker is referring to a particular taxi driver, this is not clear in (70), where the speaker may or may not have a particular taxi driver in mind. If a language has no special marker for (non-)specificity, a non-specific referent is normally introduced into the discourse through an indefinite NP (see above).³³

In the case of an NP with generic reference (+ G), we are also concerned with referential properties in that the speaker does not refer to any particular entity in the discourse world, but rather to a whole class or category of entities, as in:

(71) *Scorpions(+ G) are dangerous.*

(p.236) The idea that at some abstract level of representation the discourse operators Ω (in the NP) and Π (in the clause) essentially express the same notion (something like: determined or undetermined existence in the shared world of discourse of Speaker and Addressee) is confirmed by the fact that in certain languages the same form is used in the NP and in the clause. In Jacaltec, for example, both the exhortative mood and non-specificity are expressed by the suffix: -OJ, ‘the general suffix of irrealis’ (Craig 1977: 93).³⁴

Jacaltec (Craig 1977: 93)

(72) Way-oj ab naj
sleep-OJ EXH CLF/he
‘Would that he sleep?’

(73) X-Ø-'oc	heb ix	say-a'	hun-uj	munlabel
ASP-ABS.3-start	PL	woman	look_for-FUT	a-OJ pot
‘The women started looking for a pot’				

In (72) -OJ is suffixed to the intransitive verb in the exhortative mood, in (73) it marks the referent of the term phrase as non-specific.

Turning our attention now to *lexical discourse modifiers* in the clause (or *discourse satellites*, symbolized by Σ in the formula below), we find first of all that the occurrence or non-occurrence of an event can be emphasized by using such adverbial expressions as ‘actually’ and ‘really’.

(74) ... and then she actually insulted the waiter! (in the sense: ‘It actually happened/It was actually the case that she insulted the waiter’)

(75) (*I am telling you:*) she really went to Paris last week (‘It really happened/It was really the case that she went to Paris last week’)

Then there are Σ satellites such as ‘in vain’, as well as the so-called phasal adverbials ‘already’, ‘no longer’, ‘still’, and ‘not yet’, which are all concerned with the actuality (or non-actuality) of an event in that they express that an event ‘does or does not continue or that it has or has not come into existence’ (van der Auwera 1998a: 25; see also van Baar 1997).³⁵

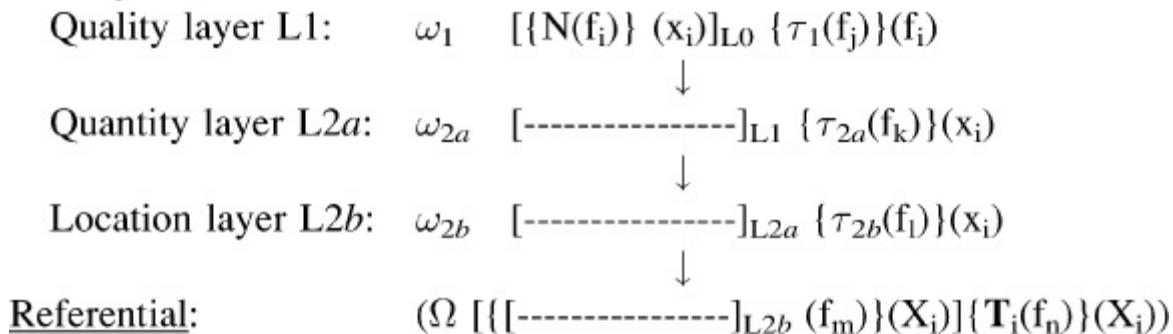
(76) *He still has his passport*

(77) *She has already left*

(p.237) It appears, then, that we can extend the parallelisms between the formal representations of the underlying structure of NP and clause (predication). For the sake of clarity I have separated the quality, quantity, and location layers:³⁶

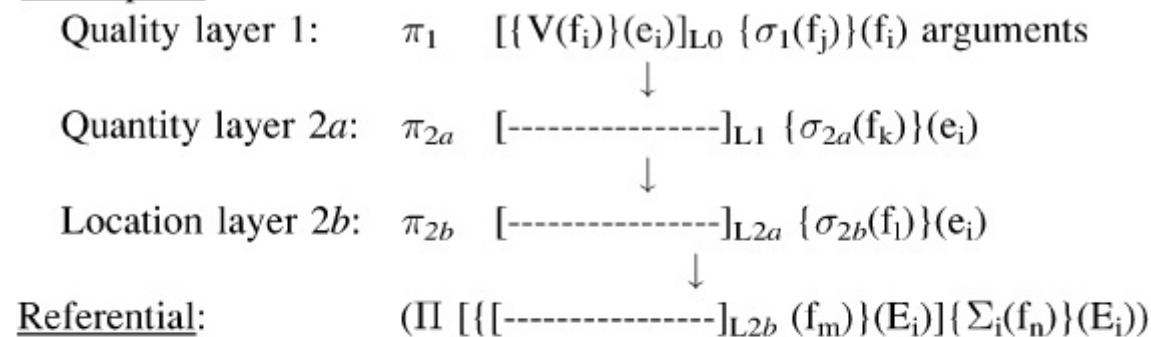
(78) UNDERLYING STRUCTURE OF THE NOUN PHRASE

Descriptive:



(78) UNDERLYING STRUCTURE OF THE CLAUSE (PREDICTION)

Descriptive:



ω_1/π_1 = qualifying NP/clause operator

ω_{2a}/π_{2a} = quantifying NP/clause operator

ω_{2b}/π_{2b} = localizing NP/clause operator

x/e = NP/clause variable

f = predicate variable

τ_1/σ_1 = qualifying NP/clause satellite

τ_{2a}/σ_{2a} = quantifying NP/clause satellite

τ_{2b}/σ_{2b} = localizing NP/clause satellite

Ω/Π = discourse operator of referent of NP/clause (predication)

X/E = referent variable of NP/clause (predication)

T/Σ = discourse satellite of referent of NP/clause (predication)

(p.238) 7.8. Conclusion

This chapter has shown that the underlying structure of the noun phrase can be represented as consisting of three hierarchically nested layers each of which accommodates both grammatical and lexical instances of the notions Quality, Quantity, and Location. Noun modifier categories relating to these three distinctions are concerned with *descriptive* properties of the referent of an NP. There are also modifiers (such as definite and indefinite articles) which specify the referential or discourse properties of a referent; these modifiers have the widest scope, extending their influence over all descriptive modifiers. I then argued that clauses (or rather that part of the clause that in FG is referred to as ‘predication’) can be analyzed in the same fashion, a descriptive part (a nucleus modified by qualifying, quantifying, and localizing operators and satellites) which is in the scope of discourse operators and satellites. This concludes the presentation of the noun and its modifier categories. The remaining chapters are concerned with the internal syntax of (integral) noun phrases.

Notes:

(1) Similarities between NPs and clauses have been observed before (cf. Van Valin and LaPolla 1997: 52 for references), but my analysis is different from the earlier proposals, because it is more comprehensive and concerns both semantic and morpho-syntactic parallels (i.e. in the underlying structure and in the actual expression).

(2) See also Rijksbaron (1989) and Dik (1997: 161).

(3) I will ignore prospective and retrospective (or perfect), because they are not pure tenses but rather mergers of tense and aspect categories. Some recent (and not so recent) contributions on deixis in general are Frei (1944), Lyons (1977: Chapter 15), Levinson (1983: Chapter 2), Greenberg (1985), Rauh (1983), Anderson and Keenan (1985), Perkins (1992).

(4) I use the term ‘event’ to refer to any type of State-of-Affairs (Action, Process, etc.; cf. Dik 1997: Chapter 5).

(5) One of the differences is that I also have a separate layer for quantifying modifiers in the underlying structure of the clause.

(6) In the layered configuration I make a basic distinction between internal and external properties. The first (inner) layer accommodates Quality operators and satellites, which relate to inherent properties of the event, action, etc. (*State-of-Affairs*) as denoted by the verb and its arguments. Quality and Location operators and satellites are both concerned with external properties and are distributed over the layers 2a (Quantity) and 2b (Location).

(7) On the relative order of adjectives, see e.g. Farsi (1968); Clark and Clark (1977: 474); Hetzron (1978b); Seiler (1978); Ney (1982, 1983); Risselada (1984); Fries (1985); Sichelschmidt (1986). For an example of stacking in a polysynthetic language (Chukchi), see Kämpfe and Volodin (1995: 102).

(8) As we saw earlier in Chapter 6, possessor phrases also relate to the location of the referent of the matrix NP by means of the so-called space-to-possession metaphor: ‘What is at one’s place is at one’s disposal’ (Claudi and Heine 1986: 316–17; Heine 1990: 144).

(9) Cf. Bartsch and Vennemann (1972: 134); Vennemann (1973: 44–5); Seiler (1978a, 1985); Moravcsik (1997: 320).

(10) The place of (in)definiteness in the underlying NP structure is discussed in section 7.6.1 below.

(11) See e.g. Lakoff and Johnson (1980: 3); Claudi and Heine (1986); Lakoff (1987); Fleischman (1989); O’Dowd (1992); Haspelmath (1996).

(12) See in this context also Gildea (1993), who shows that in Panara two synchronic tense markers are derived etymologically from demonstrative pronouns (cf. also Parker 1998, 1999, on determiners in Chamicuro).

(13) Although this is not clear from the general schema, Dik did incorporate some of the ideas proposed here and in some earlier publications (e.g. the differences between qualifying, quantifying, and localizing operators) in the second, revised edition of *The Theory of Functional Grammar*, Part 1: *The structure of the clause* (Dik 1997; see esp. Chapter 7).

(14) On pragmatic functions see also Lambrecht (1994).

(15) Hengeveld (1998) draws parallels between NPs and propositions rather than between NPs and predications. Although he also has (in) definiteness operating on the referent variable, he has identified another group of lexical modifiers as ‘restrictors’ (i.e. T satellites in my model; see section 7.7).

(16) Note that some of the symbols in my formal representations of linguistic structures are used in ‘standard’ FG with a different interpretation. For example, in FG the variable X stands for the propositional content (rather than the referent of the NP, as in my representation).

(17) Cf. Hengeveld (1992b); Keizer (1991, 1992); Fortescue (1992); Nedergaard Thomsen (1992).

(18) The idea to use the f-variable as an argument of the quality satellite in the NP is due to Kees Hengeveld (personal communication). One of the consequences of the model proposed here is that we must assume that a predicate that serves as a satellite in an NP takes a different kind of argument from a predicate that functions as the main predicate. To give an example, the argument of a predicative adjective is (the referent of) a full-blown noun phrase, whereas the argument of an attributive adjective (a level-1 satellite) is the predicate of the head noun.

(19) Cf. Langacker (1987, 1990) and Croft (1990b: 252–3) on the notion ‘physical background dimension’ or ‘base’. In cognitive grammar it is assumed that human beings only understand entities in the context of some structuring dimension of experience; the primary background dimension for nouns is space. The idea that space (but also e.g. time) is an ‘innate, universally pre-given’ category goes back at least to Kant’s *Kritik der reinen Vernunft* (published in 1781; see also Foley 1997: 84).

(20) In a way, referents can be compared with indexed records in a special file (the referent file) that is part of a language user’s knowledge base. Regarded in this way X_i ; X_j , etc. stand for different records in this referent file, each record including all sorts of information about the referent, including the source of each bit of information (who said what about X_j). Generally speaking, in the case of a definite NP this record already exists (but see the introduction to Chapter 6), whereas in the case of an indefinite NP a new record must be added to the referent file.

(21) Notice that ‘a same’ is excluded on logical grounds as the possible counterpart of ‘another’.

(22) On referring and predication as part of performing an illocutionary act, see Searle (1969: Chapter 2); cf. also Mackenzie (1987b); Dik (1997: 127 f.); Hengeveld (1998).

(23) See Dik and Hengeveld (1991) for an account of the fact that (in examples like the one given here) the complement of the verb *see* is a predication (which refers to a State-of-Affairs: what one perceives is an event, an action, etc.) and not a proposition (i.e. what one thinks, believes, fears, hopes, etc.), as in *Fred believes that Harry was talking to the police this morning*, or a speech act (what one says), as in: *Fred said: ‘Harry was talking to the police this morning.’*

(24) Since the future is a semantic category where tense and mood merge, the status of the future as a real tense has often been questioned. Part of the problem seems to lie in the fact that certain future markers are more tense-like whereas others are more modal in character. Bybee (1985: 157), for example, mentions a dialect of Yanomama as having an expected future and a desired future. Cf. also e.g. Lyons (1977: 677, 816); Ultan (1978b); Comrie (1985: 21); Chung and Timberlake (1985: 243); Palmer (1986: 216 f.).

(25) Tense is not a grammatical category in every language (Comrie 1985: 50–3; Stassen 1997: 422–87).

(26) Or as Chung and Timberlake (1985: 241) put it: ‘Whereas there is basically one way for an event to be actual, there are numerous ways that an event can be less than completely actual.’ Cf. Dik (1997: 242) and Hengeveld (1987, 1988, 1989) on the representation of mood and modality in FG; for the coding of evidential mood in languages across the world, see Chafe and Nichols (1986), Willett (1988), Johanson and Utas (1999).

Notice, incidentally, that descriptions of non-actual events may contain localizing satellites (as in ‘**Today** he may arrive in Paris’), but the fact that a non-actual event can be lexically specified for certain spatio-temporal coordinates does not give it an *actual* location in the discourse world. Such localizing modifiers in the description of the event only indicate that this event is non-actual at a certain spatio-temporal location (‘this morning’, ‘in Paris’) in the discourse world.

(27) It has been observed that in many languages the past tense has a double function in that it is also used to express unreality, normally in combination with certain other morpho-syntactic devices (conjunctions such as ‘if’ or modal adverbs such as ‘probably’). One of the explanations that has been offered to account for this relationship is that both past and unreality indicate remoteness (in time or reality), but as pointed out by Palmer (1986: 211), none of the explanations given so far is satisfactory (see e.g. Lyons 1977: 682).

(28) On the realis/irrealis distinction see e.g. Palmer (1986), Bybee et al. (1994: 236–40), in particular Elliot (2000).

(29) Of course, such modal and illocutionary distinctions can also be expressed by lexical means, e.g. modal verbs (‘I **doubt** whether she told the truth’) or illocutionary verbs (‘Then I **asked** her to leave the room’).

(30) To some extent the converse also holds: there are quite a few languages that mark the (ir)realis, but which lack tense and mood distinctions. See in this context e.g. Vidal and Klein (1998) on Pilaga and Toba; cf. also Comrie (1985: 51) on Burmese and Dyirbal.

(31) Perhaps this interesting asymmetry points to a fundamental difference between first and second order (spatial and temporal) entities (see also Harder 1996: 280). To paraphrase Chung and Timberlake (see note 26), whereas there are numerous ways that a spatial entity can be definite (see 7.6.1), there is basically one way for an event to be actual: because it (has) occurred or is occurring. Similarly: whereas there is basically one way for a spatial entity to be indefinite (see 7.6.1), there are numerous ways that a temporal entity (an event) can be less than completely actual.

(32) See Dik (1997: 176–8, 188–91) on specific and generic reference in Functional Grammar.

(33) Certain definite NPs in their so-called ‘attributive use’ (Donnellan 1966; see also Dik 1997: 189) may also be used for non-specific reference; this typically involves agentive nouns. For example, in the sentence ‘Last week Max was killed by a speeding taxi, but so far the police have not been able to identify the driver’ the fact that there is ‘a speeding taxi’ presupposes a driver; hence it is possible to talk about ‘the driver (of the speeding taxi)’ without knowing who that person is.

(34) See also Guy (1974: 29–30), who states that in Sakao nouns are inflectable for the irrealis, and Martin (1998), according to whom (ir)realis is marked by a suffix on e.g. numerals and nominals in Mocho (Maya). Du Feu (1987, 1989, 1996) discusses ‘verbal parameters expressed in the NP’ in Rapanui (Easter Island).

(35) I am grateful to Kees Hengeveld for bringing phasal adverbials to my attention.

(36) Please note that there is another important difference between the model of the underlying clause structure put forward by Dik/Hengeveld and the one that is proposed here (see also note 5). In the model that I propose (see below) all sorts of distinctions concerning the ‘actuality’ of an event are represented as operators or satellites of the referent E variable (i.e. as *discourse modifiers*), whereas in Dik/Hengeveld’s model these distinctions are represented at the same level (in the same layer) as e.g. tense.

8 Ordering Principles, Domain Integrity, and Discontinuity

8.1. Introduction

In the previous chapters we have mainly been concerned with individual constituents in the NP. After a brief survey of some general properties of NPs and their referents (Chapter 1), I argued in Chapter 2 that the various nominal subcategories define different *Seinsarten* (lexical nominal aspect) and gave many examples illustrating how the class or gender of the head noun can affect the form and order of constituents in linguistic expressions (Chapter 3). On the basis of a representative sample of some fifty languages I then gave an impression of the range of morpho-syntactic variation displayed by NP internal constituents and argued that the various modifier categories in the NP can be characterized with respect to the function that they perform in the noun phrase: some pertain to qualitative properties of the referent (qualifying modifiers; Chapter 4), other constituents specify quantitative properties (quantifying modifiers; Chapter 5), and a third category relates to locative properties of the referent (localizing modifiers; Chapter 6). Finally, in Chapter 7, I added a fourth category of noun modifier, namely discourse operators and satellites, and proposed a layered model of the NP that takes into account the differences in semantic scope between the four modifier categories distinguished in this book (*Quality*, *Quantity*, *Location*, and *Discourse*). In the same chapter I argued that the analysis of the underlying structure of the NP also applies to the clause, or rather to that part of the clause that is called *predication* in FG. Now that we have arrived at a typologically adequate model for the representation of the underlying NP structure, we can concentrate on the principles that determine the relative order of constituents in the noun phrase.

The linear organization of constituents in a linguistic expression can be seen as the consequence of many interacting ordering principles and, since these principles may counteract or reinforce each other, the linear organization in a linguistic expression will often be the result of some compromise solution (Dik 1980b, 1983a, 1986, 1997: 396).¹ Ordering principles do not necessarily have (p.240) the same impact across languages and their significance for a certain language may vary at different points in its linguistic history. Consequently, shifts in the relative impact of the various principles often cause fundamental changes in the ordering patterns of a language. These and certain other general properties of ordering principles are briefly discussed in section 8.3; for an extensive presentation I refer to Dik (1997: chapters 14–17) and Siewierska (1991: chapters 7 and 8). Section 8.4 gives a tentative typology of ordering principles; section 8.5 introduces the *Principle of Domain Integrity* (see also section 8.4.2) and discusses cases of discontinuity in the larger context of iconicity, a notion that proves to be fundamental in the explanation of word order patterns in general. First, however, I will give a more general morpho-syntactic characterization of the languages in the sample, concentrating on three major typological features: basic word order, relators (preposition, postposition, case marking), and the coding of referents in the predicate (agreement, cross-referencing).

8.2. Some morpho-syntactic properties of the languages in the sample

Especially since the publication of Greenberg's essay on language universals (Greenberg 1966a; cf. also Hawkins 1983 and Dryer 1992), it is widely known that certain morpho-syntactic properties correlate with or imply other such properties. This is discussed in greater detail in Chapters 9 and 10. Table 8.1 gives a brief typological characterization of the languages in the sample. It contains information about (1) basic word order, (2) the position of adpositions and case affixes, i.e. whether they are preposed (pr = prepositions, case prefixes) or postposed (po = postpositions, case suffixes), and (3) the morphological encoding of participants and props in the verbal complex.

Table 8.1. Some typological data of sample languages

Language	Basic word order	pr-po	Agr/XR
Abkhaz	---V/free	po	4
Alamblak	--V-	po	2
Babungo	-V--	pr	Ø
Bambara	--V-	po	Ø
Basque	---V/free	po	3
Berbice Dutch Creole	-V--	pr po	Ø
Bukiyip	--V-	po	2
Burmese	---V	po	Ø
Burushaski	---V	po	2
Cayuga	(see below)		
Chukchi	-VV-/free	po	2
Dutch	-V--/-V-	pr po	1
Galela	--V-	pr po	2
Georgian	--V-/free	po	2 ¹
Greenlandic, West	--V-	po	2
Guaraní	-V--/free	po	1
Gude	VV--	pr	Ø
Hittite	---V?	po	1
Hixkaryana	-V--	po	2
Hmong Njua	-V--	pr	Ø
Hungarian	--V-/free	po	2
Hurrian	---V?	po	1
Ika	---V	po	2
Kayardild	--V-/free	po	Ø
Ket	-VV-	po	3

Kisi	-V--/-V-	pr po	Ø
Koasati	--V-/free	po	4
Korean	---V	po	Ø
Krongo	V---	pr po	1
Lango	-V--	pr	2
Mandarin Chinese	--V-/free	pr po	Ø
Nahali	---V?	po	?
Nama Hottentot	--V-	po	Ø
Nasioi	--V-?	po	2?
Ngalakan	--V-	po	2
Ngiti	-V--	po	1
Nivkh	---V?	po	1 (see below)
Nung	-V--	pr	Ø
Nunggubuyu	-VV-/free	po	2
Oromo	---V	pr po	1
Pipil	V---	pr	2
Quechua, Imbabura	--V-?	po	1
Samoan	V---	pr	1
Sarcee	--V-	po	2
Sumerian	---V	po	2
Tamil	---V?	po	1
Tsou	V---	pr	Ø
Turkish	--V-	po	1
Vietnamese	-V--	pr	Ø
Wambon	--V-	po	1

(1) As a rule the Georgian verbal complex contains a prefix that agrees ‘in person/number with the (in)direct object if it is at least a first or second person pronoun and with the subject otherwise; the suffix agrees in person/number normally with the subject or under special conditions with the object’ (Testelec 1997a: 236; cf. also Fähnrich 1986: 72).

(p.241)

(p.242) As regards basic word order, I have classified languages in terms of the conventional types: *V-initial* (V---), *V-second* (-V--), *non-rigid V-final* (--V-), and *rigid V-final* (---V), thus temporarily ignoring the fact that languages such as Samoan do not have a *distinct* class of verbs (see 1.5.1). Other entries in Table 8.1 (e.g. -VV-) and question marks are explained below. The word order patterns in Table 8.1 only relate to the position of the main verb in the main clause. Dutch and Kisi have two entries: -V-- and --V-. In Dutch the finite (main or auxiliary) verb normally occupies the second position in the main clause, whereas the non-finite verb follows the object:

Dutch

(1) Jan kocht een boek [SV_{fin}O]

Jan bought a book

'Jan bought a book'

(2) Jan heeft net een boek gekocht [SAux_{fin}OV]

Jan has just a book bought

'Jan has just bought a book'

Kisi has a similar kind of alternation. Clauses with ‘simple verbs’ have SVO order, whereas clauses with a ‘compound verb’ (i.e. involving an auxiliary and a non-finite verb) are non-rigid V-final with Aux in second position.

Kisi (Childs 1995: 249–50)

(3) kèùwó lòwá sàá

snake bite Saa

'The snake bit Saa'

(4) Fàlà có lééñndó yìkpàá

Fala AUX machete sharpen

'Fallah is sharpening the machete'

The third column indicates whether there is an agreement (Agr) or a cross-referencing (XR) relation between one or more terms and one or more pronominal elements in the verbal complex (more on the difference between agreement and cross-referencing in section 8.2.3). The cardinal number indicates the maximum number of arguments that can be coded simultaneously in the predicate. However, since the sample contains only a single language with a genuine agreement relation between the predicate and its arguments (namely Dutch), in nearly all languages the cardinal number stands for the maximum number of NPs that can be cross-referenced in the predicate.² Finally, ‘Ø’ means that there is neither agreement nor cross-referencing and a question mark means that the information is lacking or uncertain.

(p.243) 8.2.1. Word order

Since the notions subject and object are only judged to be relevant in a subset of the world’s languages (Dik 1997: chapters 10–11) it is not feasible to classify languages in terms of the relative order of S(subject), O(object), and V(verb). Instead I have attempted to classify languages on the basis of the position that the main verbal predicate occupies in the basic pattern, which can be defined as the most frequent pattern with two nominal arguments.³ Here are examples of the four types:

Tsou (Szakos 1994: 168): V-INITIAL (V---)

(5) m-o	mo-oeai	no	emoo	'e	avai
ActFoc-PRET	ActFoc-make	ART(PROX)	house	ART	Avai
'Avai builds houses' ⁴					

Lango (Noonan 1992: 121): V-SECOND (-V-)

(6) lócé òmìiyì	àtîn	búk	bót	dákô
man	3SG:give:BEN.PERF	child	book	to woman
'The man gave the book to the woman for the child'				

Nama Hottentot (Hagman 1974: 149, 212): NON-RIGID V-FINAL (--V-), i.e. certain non-complex sentence constituents may or must follow the verb (see Hagman 1974: 110 ff. on the subordinative suffix (SubSx) à): no constituents after the verbal complex

(7) 'áo-p	ke	tará-s-à	péré-p-à
man-3SG.M	DECL	woman-3SG.F-SubSx	bread-3SG.M-SubSx
kè	màa		
RemP	give		
'The man gave the woman bread'			

with postpositional phrase following the verb⁵

(8) siíke ke	//na~ápá	kè	síí	paa-s	/xàa
we	DECL	there	RemP	arrive	bus-3SG.F by
'We arrived there by bus'					

Korean (Kim 1987: 894): V-FINAL (---V)

(9) John i	Mary	eke	c ^h æk	üł	cuəs'ta
John	NOM	Mary	DAT	book	ACC gave ⁶
'John gave a book to Mary'					

(p.244) Additionally a language may be classified as 'free' (cf. Dryer 1989c). Although in the functional paradigm free word order languages are not deemed to exist (Dik 1997: 394), I will use this as a label to indicate that word order is determined by pragmatic rather than syntactic or semantic considerations. An example of such a language is Abkhaz (Hewitt 1979: 51; see also p. 103): 'The most frequent word-order ... is S-IO-DO-V, although any order is theoretically possible. The referent of any other intra-verbal affix will most naturally stand between IO and DO'. More specifically, in this context, 'free' mainly relates to the position of the verb. That is, a V-final language in which NPs can occur in any possible order as long as the verb remains in its sentence-final position is still put down as ---V, not as 'free'.⁷ If, however, there is

considerable variation in the position of the verb, the language is (also) classified as ‘free’ in Table 8.1 above; this holds for e.g. Basque (Saltarelli 1988: 66–7), Chinese (Li and Thompson 1989: 19), Guaraní (Gregores and Suárez 1967: 182), Hungarian (de Groot 1989: 23), Kayardild (Evans 1995: 1, 92), and Koasati (Kimball 1991: 513–14).

Sometimes it was impossible to give one basic order, as when different ordering patterns seemed equally frequent. This seems to hold for the following languages: Chukchi (see below), Gude (Hoskison 1983: 106; speakers of Gude use V-initial order with some verbal aspects and V-second orders with others), Ket, and Nunggubuyu (Heath 1984: 509 f.). Consider, for instance, Comrie’s characterization of Ket (Comrie 1981a: 265; cf. also Werner 1997a: 332 f.):

In terms of word order typology, Ket has many of the characteristics of the subject-object-verb type, although the order of major constituents within the clause varies between subject-verb-object and subject-object-verb; given the prolific system of subject and object agreement [= cross-referencing; see below—JR], many transitive sentences do not have their full complement of noun phrase arguments.

Often such languages also have more or less ‘free’ word order (Chukchi, Nunggubuyu). Consider, for instance, these remarks about Chukchi (Comrie 1981a: 251): ‘[Where separate words rather than incorporation are used], Chukchi word order is very free: while subjects usually precede their verb, the predominance of object-verb over verb-object is slight.’ The rough division of fifty-two languages into just four word order types is not unproblematic. It is evident from these quotations that, to the extent that this is at all possible, I have abstracted away from the fact that certain languages do not lend themselves easily for a classification in terms of basic word order (cf.

Dryer 1989c, 1997; also Mithun 1992). This may be due to a variety of reasons, which also tend to conspire. To name a few: not every language seems to have a basic word order, certain languages make extensive use of just verbal complexes (i.e. there is hardly any syntax, but all the more (p.245) morphology), in some languages clauses or NPs consist of appositional and often discontinuous constituents.⁸

Last but not least, many authors have remarked on the fact that it is often rather unusual and sometimes virtually impossible to have two full (definite) NPs in one sentence (e.g. Heath 1984: 507, 513; Edwards 1985: 398; Siewierska 1988: 9; Bybee et al. 1990: 5; Barnes 1990: 289–90; MacDonald 1990: 4; Mithun 1992).

For a number of V-final languages it could not be determined whether the language in question belongs to the rigid or to the non-rigid subtype.⁹ This is indicated by the question mark: Hittite (Tomlin 1986: 188; but cf. Garrett 1990: 274), Human, Nahali, Nasioi, Nivkh (Comrie 1981a: 270; Gruzdeva 1998: 40), Imbabura Quechua (Cole 1982: 72), and Tamil (Asher 1982: 25; Hock 1989: 400). Often this is due to the fact that there were no explicit statements or clear examples in the grammar of the language in question, and sometimes it is due to the fact that it was not possible to decide whether some postverbal constituent was part of the sentence proper, or that it followed the verb as an afterthought, thus occurring outside the real sentence.

8.2.2. Relators¹⁰

The second column refers to the position of relators: adpositions and case affixes. Some languages have both preposed and postposed relators: Dutch, Mandarin Chinese (Li and Thompson 1989: 24–5), Berbice Dutch Creole (Kouwenberg 1991: 135 ff.),¹¹ Galela (van Baarda 1908: 153), Kisi (Childs 1995: 129), Krongo (which has proclitic case markers and postpositions; Reh 1983), and Oromo (Bender et al. 1976b: 139).¹²

Of course, these are not the only means languages use to express a semantic, syntactic, or pragmatic relation, but the other possibilities (e.g. preverbs (Abkhaz), serialization (Berbice Dutch Creole), coverbs (Chrau), tone (Babungo), or word order (Asmat)) are irrelevant in this context.

(p.246) 8.2.3. The coding of (properties of) referents in the predicate

Basically there are three types of languages with respect to the way properties of referents are coded in the predicate: (1) agreement languages, (2) cross-referencing languages, (3) languages in which no properties of referents are coded in the predicate.

The difference between cross-referencing and agreement is that in the first case the predicate plus ‘incorporated’ pronominal elements alone can (and often will) function as a complete sentence by itself; if necessary (e.g. for reasons of emphasis, or to avoid possible ambiguity) full terms or free personal pronouns can be added (e.g. in pseudo-English: ‘[The man/he] he-gave-her-it [the book/it] [to his wife/to her]’). That is, distinct terms denoting the same referents as those denoted by the pronominal elements in the predicate may occur in the same sentence, but they are basically redundant. In an agreement relation, on the other hand, the verbal complex cannot constitute a complete sentence by itself, although it does formally adjust to (‘agree with’) one or more properties of the referent(s), as in the Dutch examples below.¹³ The maximum number of referents that can be cross-referenced in the verbal complex is four. Consider this example from Abkhaz (Hewitt 1979: 51):

Abkhaz (Hewitt 1979: 51)

(10)	a-xàc' a	a-pħ°ès	l-y°èza	a-š°q°è
	the-man	the-woman	her-friend	the-book
(ϕ-)lə + z-lə-y-te-yt'				
(it-)her + for-to_her-he-give-FIN				
'The man gave the book to the woman for her friend'				

Nivkh is unusual in that the category person (and often number as well) is only marked on non-matrix verbs (Mattissen and Drossard 1998: 24; see also Gruzdeva 1998: 48, 55); -t is the first person converbal suffix):

Nivkh (Mattissen and Drossard 1998: 24; from Panfilov 1962–5: ii. 147):

(11)	ñi	meɻtu-řo-t	ŋa-ŋəŋ-t	vi-đ
	1SG	rifle-take-CV.1	beast-hunt-CV.1	go-IND/NML

'I take my rifle and go hunting'

I have already mentioned that the sample contains only a single clear example of a language that has an *agreement relation* between lexical NPs and the verb, namely Dutch, in which the verb agrees in number/person with the nominal or pronominal subject NP:

Dutch

- (12) Ik loop
1SG.S walk

(p.247)

- (13) jij loop-t (but: loop jij? [walk you] 'Are you walking?')
2SG.S walk-PRES.2/3SG

- (14) hij/ de jongen loop-t
3SG.M.S/ the boy walk-PRES.2/3SG

- (15) wij/jullie loop-en (spelled: lopen)
1PL.S/2PL.S walk-PRES.PL

- (16) ze/ de meisje-s loop-en (spelled: lopen)
3PL.S/ the girl-PL walk-PRES.PL

It seems that the agreement relation, which is so well attested in languages spoken on the European continent, is a very marked type from a cross-linguistic perspective.¹⁴

Languages of the third type (no marking) seem to occur mostly on the African continent (Babungo, Bambara, Gude, Kisi, Nama Hottentot) and in South-East Asia (Burmese, Mandarin Chinese, Korean, Hmong Njua, Nung, Tsou, Vietnamese).

Babungo (Schaub 1985: 38)

- (17) Lāmbí sāŋ ɳwé bù' lāa ɳwé shàw fá
Lambi beat:PERF him because that he steal:PERF thing
'Lambi has beaten him because he has stolen something'

Hmong Njua (Harriehausen 1990: 55)

(18) Kuv	yuav	lug	tsev	pig kig
1SG	FUT	come	house	tomorrow
'I will come home tomorrow'				

8.3. General properties of ordering principles

Ordering principles do not *directly* determine the order of constituents in a language. In FG a distinction is drawn between (i) *ordering principles*, with which we will be concerned here and which are potentially relevant to all languages, and (ii) *placement rules*, which belong to the grammars of individual languages. Ordering principles may be regarded as forces that together define the working space of the language specific placement rules. In other words, ordering patterns determined by placement rules must be in accordance with universal ordering principles. For instance, the *Principle of Scope* (Chapter 10) says that in the simple, integral NP the demonstrative (dem) should precede the numeral (num) if both occur before the head noun (N) and that it should follow the numeral if both follow the head noun. This yields four possible orderings: [dem num N], (p.248) [dem N num], [num N dem], and [N num dem]. Which of these patterns a language uses is irrelevant at this level of abstraction. Whether the demonstrative or the numeral precedes or follows the noun in a particular language is determined by placement rules, which belong to the grammars of the individual languages.

Ordering principles and placement rules are part of the *expression component* in the FG model, which is also responsible for the *form* of the constituents (inflectional morphology) and the *prosodic contour* of the expression. The rules and principles that are part of the expression component apply to fully specified underlying structures (Chapter 7). Since the underlying structure is a semantic rather than a syntactic structure, there is no need for transformations (in fact they are not permitted in FG); constituents are immediately placed in the right position.

I have already mentioned above that ordering principles do not necessarily have the same impact across languages, but so far this has not been investigated in any detail. The fact that some principles may be more relevant than others in a certain language is often related to the trade-off between syntax and morphology, both of which concern the formal expression of the underlying structure. The more of the underlying structure can be expressed through morphological means, the less need be taken care of by syntax and vice versa (Dik 1997: 392). For instance, in languages where the verbal complex alone may constitute a complete sentence (as in e.g. Abkhaz and West Greenlandic), the role of syntax will be relatively small as compared to languages of the isolating type (e.g. Chinese and Nung), in which the role of morphology is greatly reduced and the burden of expression is almost entirely delegated to the syntactic rules of the language.

8.4. A typology of ordering principles

Word order rules and principles can be divided according to the *reference point* relative to which the position of a constituent is defined.¹⁵ There are three kinds of reference points: (i) domains (a domain can be defined as the head and its dependants, e.g. clauses and NPs), (ii) boundaries of domains (e.g. clause-initial, clause-second, clause-

final position), and (iii) constituents in domains (in particular head constituents such as verbs and nouns).¹⁶ Here are some examples.

(p.249) 8.4.1. Type I: placement relative to a domain

An example of a word order principle that defines the position of a constituent in a linguistic expression relative to a *domain* (in this case: the domain is an NP) is the principle which deals with the position of the Subject NP, which is defined relative to the Object NP (Dik 1997: 405):

- (19) *The Subject position precedes the Object position.*

Actually this principle is a specification of a more general ordering principle, the *Principle of Functional Stability*, which says that constituents with the same functional specification are preferably placed in the same position relative to each other.

Another example is the *Principle of Increasing Complexity*, which, as Dik (1997: 404) notes, was first formulated by Behaghel (1932) as the *Gesetz der Wachsenden Glieder* ('law of increasing parts'):

- (20) There is a preference for ordering constituents in an order of increasing complexity.

The same principle has a more specific version, known as the *LIPOC principle*, where LIPOC is short for 'language independent preferred order of constituents' (*ibid.* 411):

- (21) Other things being equal, constituents prefer to be placed in an order of increasing complexity, where complexity of constituents is defined as follows:

- (i) clitic < pronoun < NP < adpositional phrase < subordinate clause;
- (ii) for any category X: X < X coordinator X;
- (iii) for any categories X and Y: X < X subordinator Y.

Principles (19) and (21) may define conflicting preferences, as in the case of the Dutch subject clause *Dat ze niet kwam* 'That she did not come' in the following example:

Dutch

(22)	Dat	ze	niet	kwam	verbaas-de	me
	that	3SG.F.S	not	come:SG.PAST	surprise-SG.PAST	1SG.O
'That she did not come surprised me'						

Observe that the subject appears before the object, as it should according to (19). Due to LIPOC the following (preferred) construction is also attested; now the 'heavy' subject appears last, after the less complex pronominal object, and a dummy pronoun (*het* 'it') occurs in clause-initial position:

Dutch

(23)	Het	verbaas-de	me	dat	ze	niet	kwam
------	-----	------------	----	-----	----	------	------

it surprise-SG.PAST 1SG.O that 3SG.F.S not come:SG.PAST

'It surprised me that she did not come'

(p.250) 8.4.2. Type II: placement relative to the boundary of a domain

Several ordering principles relate to the boundary of a domain. The *Principle of Domain Integrity* (Rijkhoff 1990a), for instance, states:

- (24) Constituents prefer to remain within the boundaries of their domain; constituents of a domain prefer not to be interrupted by embedded domains.

The *Principle of Domain Integrity* accounts for restrictions against discontinuous constituents. Exceptions to this principle are attributable to more pragmatically motivated ordering principles, such as the *Principle of Pragmatic Highlighting*. According to this principle, constituents with special pragmatic functionality (New Topic, Given Topic, Completive Focus, Contrastive Focus; Dik 1997: chapter 13) are preferably placed in special positions, which always include the clause-initial position. A more specific instance of this principle reads (Dik 1997: 408):

- (25) There is a universally relevant clause-initial position P1, used for special purposes, including the placement of constituents with Topic or Focus function.

The following example from Dutch (Geerts et al. 1984: 962) illustrates how an embedded domain (*van Walschap* 'by Walschap', which has contrastive focus here) may be expressed discontinuously from its matrix domain (*twee romans* 'two novels') under the influence of the *Principle of Pragmatic Highlighting*:

Dutch

(26) Van Walschap heb ik twee romans gelezen

by Walschap have:1SG.PRES 1.SG.S two novels read: PastPart

'I have read two novels by Walschap'

In the next example only the head (*kinderen* 'children'; like *Walschap* in the previous example to be pronounced with emphatic stress) is placed in clause-initial position (ibid. 961):¹⁷

Dutch

(27) Kinderen had hij op straat vrij veel gezien

children have:SG.PAST 3M.SG.Son street rather many see:PastPart

'He had seen quite a few children in the street'

(p.251) 8.4.3. Type III: placement relative to a constituent in a domain

An example of a principle that uses a constituent in a domain as its reference point is the *Principle of Head Proximity* (Rijkhoff 1986), which is discussed in more detail in Chapter 9. The short version of it reads (Dik 1997: 402):

- (28) Constituent ordering rules conspire to keep the heads of different domains as close together as possible.

Another example of a Type III principle is the *Principle of Centripetal Organization* (*ibid.* 401):

- (29) Constituents conform to [this principle] when their ordering is determined by their relative distance from the head, which may lead to ‘mirror image’ ordering around the head.

Furthermore, all principles that refer to the *Prefield* or the *Postfield* (i.e. a position in the area before or after the head in a domain, respectively) are also of Type III.

8.4.4. The mixed type: principles with different kinds of reference points

If different kinds of reference points are used in the definition of an ordering principle, this seems to be due to one of two factors:

- (a) the constituents under consideration form a disjunctive category;
- (b) the ordering principle in question is based on various more general ordering principles.

The *Relator Principle*, which concerns the placement of adpositions, case markers, and conjunctions, is an example of (a). In this principle the *relata* serve as reference points:

- (30) Relators have their preferred position

- (i) in between their two relata;
- (ii) at the periphery of the relatum with which they form one constituent (if they do so).

The category of relators is in reality a mixed bag, consisting of different kinds of function words, which serve to join (and to mark a particular relationship between) relata. For example, adpositions and case markers can be used to express a relationship between an argument NP (a domain) and the main predicate (a head). Only coordinating conjunctions indicate a relationship between entities of a kind, such as clauses or NPs. In sum, the reference points implied in the *Relator Principle* are of a different kind (domains as well as heads), because the category of relators consists of different kinds of parts-of-speech.

The principle that deals with the position of grammatical tense-aspect-mood markers or π -operators (Dik 1997: 413; see also Chapter 7) is an example of a (p.252) principle that is the consequence of the combined functioning of several more general principles. According to this principle π -operators prefer expression:

- (a) in the counterfield (the counterfield is the Postfield in a Prefield or head-final language and the Prefield in a Postfield or head-initial language; obviously the head is used as a reference point here), or
- (b) the second position in the clause (now the boundary of the domain is used as a reference point).

Placement in the counterfield, however, can be attributed to the *Principle of Head Proximity* (see also above); otherwise the operators would occur in between V and the heads of the Subject and Object NP:¹⁸

PREFIELD LANGUAGES

preferred: P1 S O V π -operators [π in counterfield]

non-preferred: P1 S O π -operators V

POSTFIELD LANGUAGES

preferred: P1 π -operators V S O [π in counterfield]

non-preferred: P1 V π -operators S O

Placement in the second position of the clause is due to a combination of three principles. First of all, in many so-called Prefield languages certain phrases may or must follow the verb. Should this happen, the π -operators would occur in between heads of domains V and N, which would be at variance with the *Principle of Head Proximity*:

NON-RIGID PREFIELD LANGUAGES

non-preferred: P1 S O V π -operators X Y counterfield

The situation may be resolved if the π -operators occur elsewhere, and as they are relatively ‘light’ constituents in terms of internal complexity (compared to NPs), we may expect them to occur early in the clause according to the *Principle of Increasing Complexity*. They are not preferred somewhere in between the Subject and Object NPs because of the *Principle of Head Proximity*, and not in P1 either, which is reserved for pragmatically highlighted constituents. This leaves us with the second position:

NON-RIGID PREFIELD LANGUAGES

preferred: P1 π -operators S O V X Y 2nd position

Thus, the occurrence of π -operators in clause-second position is the result of the combined functioning of the *Principle of Head Proximity*, the *Principle of Increasing Complexity*, and the *Principle of Pragmatic Highlighting*.

(p.253) 8.5. Iconicity, domain integrity, and discontinuity

Iconicity is a cover term for all those instances in which the form or structure of a linguistic expression reflects in one way or another our conceptualization of forms or structures in the external, physical world. Although iconicity manifests itself in many

different ways, we will only be concerned with one particular form: *structural isomorphism* (Croft 1990a: 174), i.e. the correspondence between the linear organization of a linguistic utterance and the underlying structure of that utterance. Croft (1990a: 171–2) points out that iconicity of human language is difficult to prove independently, because most of what we know about it is based on linguistic evidence. Consequently, our recognition of iconicity as a major force in constituent ordering may in fact be a self-fulfilling hypothesis. In our case, however, this problem is not urgent, because we do not base ourselves on utterances (phonological, morphological, and syntactic structures), but on underlying structures, which are *independently* motivated semantic representations.

Iconicity is probably one of the most pervasive forces that determines the actual form and order of constituents in an utterance, as can be illustrated by the fact that both the *Principle of Domain Integrity* and the ordering principles that will be discussed in the following chapters are in reality all elaborations of a more general iconic principle already formulated by Behaghel (1932: 4):¹⁹ ‘Das oberste Gesetz ist dieses, daß das geistig eng Zusammengehörige auch eng zusammengestellt wird’ [The principal law is this: that what belongs together semantically is also placed together syntactically]. The tendency to preserve semantic aspects (as more or less reflected in the underlying structure) in linear order is usually explained as a consequence of the way we conceptualize relations between entities. It is assumed that semantically related elements are conceptualized in terms of spatial distances: the stronger the semantic relation between elements (or the closer the semantic bond), the shorter the distance between these elements (cf. Lakoff and Johnson 1980). The *Principle of Domain Integrity* (a Type II principle; see above) is perhaps the most direct reflection of Behaghel’s iconic principle and can be regarded as the FG equivalent of ‘constituency’ in other theories. The *Principle of Domain Integrity* accounts for the fact that discontinuity is a marked phenomenon:

Constituents prefer to remain within the boundaries of their proper domain; constituents of matrix domains prefer not to be interrupted by embedded domains.

(p.254) Not surprisingly, the strong tendency of words to remain within their domain has also been explained in the light of language production and comprehension. Words that form a semantic unit are expected to be processed faster and easier when they occur together, and in speaking or writing this necessarily means: in an uninterrupted, continuous sequence. This claim can be substantiated by a body of psycholinguistic evidence (e.g. Bever 1970; Lindsley 1975; Grosjean et al. 1979). There is also evidence from language acquisition to support the *Principle of Domain Integrity*. In his work on the universal operating principles children employ in language acquisition, Slobin (1985: 228) proposed the following non-interruption principle: ‘Avoid interruption or rearrangement of linguistic units.’

The *Principle of Domain Integrity* reduces the number of logically possible NP-internal word order patterns in two ways. In the first place it states that constituents of the NP are normally expressed in one continuous string (notice that this also applies to constituents of embedded domains in the NP, such as possessor NPs and relative clauses). Secondly the *Principle of Domain Integrity* states that the preferred position of embedded domains is in the periphery of the matrix domain. In this section we will only be concerned with the first of the two propositions that this principle contains; the

second proposition is discussed in connection with the *Principle of Head Proximity* and the *Principle of Scope* (see Chapters 9 and 10, respectively).

If the demonstrative, the numeral, the adjective, and the noun should all occur in the same (integral) NP, this gives us $4 \times 3 \times 2 \times 1 = 24$ possible orderings.²⁰

(31) Possible patterns in count noun languages: Principle of Domain Integrity

dem num A N	dem num N A	dem N num A	N dem num A
num dem A N	num dem N A	num N dem A	N num dem A
A dem num N	A dem N num	A N dem num	N A dem num
dem A num N	dem A N num	dem N A num	N dem A num
num A dem N	num A N dem	num N A dem	N num A dem
A num dem N	A num N dem	A N num dem	N A num dem

Since numeral classifier languages can have one more constituent in the simple NP (namely the numeral classifier), the number of patterns in these languages would be five times the number calculated for noun languages that do not require a numeral classifier, i.e. $5 \times 4 \times 3 \times 2 \times 1 = 120$ logically possible patterns. However, the numeral and the classifier always form an inseparable combination (Allan 1977; Greenberg 1975), so that the total number of possible NP patterns (p.255) amounts to forty-eight, twice the number of possible patterns calculated for count noun languages.

(32) Possible patterns in numeral classifier languages: Principle of Domain Integrity

dem num clf A N	dem num clf N A	dem N num clf A	N dem num clf A
num clf dem A N	num clf dem N A	num clf N dem A	N num clf dem A
A dem num clf N	A dem N num clf	A N dem num clf	N A dem num clf
dem A num clf N	dem A N num clf	dem N A num clf	N dem A num clf
num clf A dem N	num clf A N dem	num clf N A dem	N num clf A dem
A num clf dem N	A num clf N dem	A N num clf dem	N A num clf dem
dem clf num A N	dem clf num N A	dem N clf num A	N dem clf num A
clf num dem A N	clf num dem N A	clf num N dem A	N clf num dem A
A dem clf num N	A dem N clf num	A N dem clf num	N A dem clf num
dem A clf num N	dem A N clf num	dem N A clf num	N dem A clf num
clf num A dem N	clf num A N dem	clf num N A dem	N clf num A dem
A clf num dem N	A clf num N dem	A N clf num dem	N A clf num dem

The fact that in numeral classifier languages the numeral concatenates with a classifier to form a classifier phrase is explained by Greenberg (1975), who convincingly argues that the numeral + classifier combination was originally an adverbial quantifying phrase, which, in due course, has become integrated in the NP.²¹ Hence one may

assume that ultimately the inseparability of the numeral and the classifier can also be regarded as a consequence of the *Principle of Domain Integrity*.²²

8.5.1. Counter-examples to the Principle of Domain Integrity: discontinuity

It is important to distinguish between real and apparent cases of discontinuity. Sometimes a constituent may appear to occur outside its proper domain, whereas strictly speaking this is not the case, because it is not part of an *integral NP*.²³ Instead the constituent in question (i) is in apposition (section 1.5.4.2), or (ii) is itself a (non-apposed) distinct phrase at clause level.

8.5.1.1. Apparent discontinuity

Consider again, for example, these sentences from Kalkatungu (Blake 1983: 145; see also section 1.5.4.1), which is not in the sample; the apposed constituents ([p.256](#)) are in bold print (for an extended discussion on apposition, see Matthews 1981: 224–36).

Kalkatungu (Blake 1983: 145)

- | | | | | | |
|------------------|-----------------|---------------|---------------|---------------|------|
| (33) a. cipa-yi | <u>t</u> uku-yu | yaun-tu | yani | icayi | |
| | this-ERG | dog-ERG | big-ERG | white_man | bite |
| b. cipayi | <u>t</u> kuyu | yani | icayi | yauntu | |
| c. <u>t</u> kuyu | cipayi | icayi | yani | yauntu | |
| d. yauntu | cipayi | <u>t</u> kuyu | icayi | yani | |
| e. cipayi | icayi | yani | <u>t</u> kuyu | yauntu | |
| f. yani | icayi | cipayi | yauntu | <u>t</u> kuyu | |
- 'this big dog bit/bites the white man'

Appositional NPs are attested in quite a few of the Australian languages. In the sample Ngalakan and Nunggubuyu are believed to lack integral NPs.²⁴ Merlan (1983: 83) writes:

In general, the Ngalakan NP exhibits a fairly loose sort of structure. It is possible for constituents of what could be considered the ‘same’ NP to be separated from each other by other clausal constituents, or for many NP constituents having the same referent to be strung together in a fairly loose sort of appositional structure.

According to Heath (1984: 505–6) one cannot seriously speak of phrasal NPs in Nunggubuyu either:

My general conclusion is that, while it is meaningful to group coreferential words together in some way, being associated in each ‘clause’ with a particular grammatical relation ..., these words commonly occur in surface structure either separated by pauses (or intonational breaks) or by intervening constituents. ... Personal pronouns, demonstratives, ‘adjectives’ ..., and relative clauses (including genitive usage), may all occur independently as one word units; when they do occur in the same ‘clause’ as a (nuclear) noun, it is possible to regard the combinations as appositional rather than phrasal, but a fuzzy NP unit might be

recognized in cases where the sequence is pronounced rapidly and without pauses or intonational boundaries.

It may also be the case that what seems to be a noun modifier is in fact an adverb(ial) constituent at the level of the clause. A case in point is numerals in Hixkaryana, which (as was already mentioned in section 1.5.4.1) Derbyshire (1979: 103) categorized as sentence adverbs.

To sum up, in cases of apparent discontinuity the constituent in question is not part of an integral NP. Instead, it can be a referring expression that has an appositional relationship with another NP, as in e.g. the Australian languages ([p.257](#)) referred to above; or it should be regarded as a distinct (i.e. non-apposed) clausal constituent by itself, as in e.g. Hixkaryana (cf. Siewierska 1988: 166–70).

8.5.1.2. True discontinuity

Let us now turn to real instances of discontinuity, whereby a constituent of an integral NP occurs outside its proper domain. Although the evidence unfortunately is often very fragmentary, such instances of true discontinuity can typically be attributed to some pragmatic factor, complexity, or both. In Polish everyday speech, for instance, discontinuity is only regularly used in certain contrastive contexts, as in (Siewierska 1984; cf. also Siewierska and Uhlířová 1997: 137):

Polish (Siewierska 1984: 59)

(34) A:	Słyszałam,	że	dostałaś	złotą
	heard:PAST.1SG	that	give:PAST.2SG.F	gold:ACC.SG.F
	brązoletkę.			
	bracelet:ACC.SG			
	'I heard you got a gold bracelet'			
B:	Znowu	ci	się	coś
	again	2SG.DAT	REFL	something:NOM
	pomyliłam			
	be_mistaken:PAST.3SG.N			
	'You've heard it wrong'			
	Brązoletkę	dostałam	srebrną.	
	bracelet:ACC.SG	give:PAST.1SG.F	silver:ACC.SG.F	
	'I got a silver bracelet'			
	Złoty	dostałam	pierścionek.	
	gold:ACC.SG.M	give:PAST.1SG.F	ring:ACC.SG	
	'I did get something gold, but it was a ring'			

In this example ‘gold’ is contrasted with ‘silver’ and ‘bracelet’ with ‘ring’ (notice that ‘gold’ and ‘silver’ agree in gender, number, and case with the head noun). Siewierska

(1984: 62–4) found that sentences containing discontinuous constituents share a number of characteristics in Polish:

- (a) Discontinuous phrases typically occur in clauses where the subject is expressed solely by the agreement suffixes on the verb.
- (b) Discontinuity within the NP tends to involve phrases containing a noun and just one modifying adjective, possessive pronoun or deictic. Phrases with sequences of modifiers are very rarely discontinuous.
- (c) Discontinuity within the PP [prepositional phrase] can only occur if the prepositional object is modified by an adjective, possessive pronoun or deictic. Moreover, no part of the prepositional object may precede the preposition ... In addition, PPs within the NP cannot be discontinuous, since PPs as units are not marked whether (p.258) they occur in NP, VP or S, and thus if moved could not be unambiguously associated with the head of the NP ...
- (d) The constituents of the phrase are split up into two parts such that one part occurs preverbally, in speech always initially, while the other is placed in final position. Hence they must be separated by the verb
- (e) Typically only pronouns, clitics, adverbials and less often short PPs can intervene between the verb and final constituents. Clauses with direct and indirect objects in this position are unacceptable: ...
- (f) Only one instance of discontinuity per clause is tolerated. This is a consequence of (a) and (e).

Since one may assume that sentences containing discontinuous constituents are relatively difficult to process, these properties can be explained as an attempt to avoid making the sentence more complicated than is strictly necessary for the hearer to interpret.

Nearly all cases of discontinuity reported in the grammars of sample languages involve embedded domains (but see Vilkuna 1997: 223 f. on discontinuity of demonstratives and other noun modifiers in the Uralic languages).²⁵ Although this study is mainly about simple NPs, here are some examples.

We have seen above that discontinuity in Polish is primarily associated with contrastive contexts. Similarly, only ‘counter-assertive’ numerals in Babungo are also expressed discontinuously. Counter-assertive numerals of the subject NP are placed in clause-final position:

Babungo (Schaub 1985: 123)

(35) a.	mè yè	nyáasə	sèbòò	táa	yìb̥i i	yô
	I	see:PERF	animals	two	in	pit that
‘I saw two animals in that pit’						
b.	mè yè	nyáasə	táa	yìb̥i i	yô	sèbòò
	I	see:PERF	animals	in	pit that	two
‘I saw <u>two</u> animals in that pit’						

We have already seen that in Ngiti highly emphasized numerals may appear in clause-initial position:

Ngiti (Kutsch Lojenga 1994: 355)

(36) áru	ma	m-òkèrè	itsu
eight	1SG	SC-cut:PERF.PRES	tree
'I have cut eight trees'			

(p.259) In Modern Greek, which is not in the sample, a genitive NP (in bold) can occur in NP-initial or in clause-initial position for emphasis; compare:

Modern Greek (Joseph and Philippaki-Warburton 1987: 107)

(37) a.	θavmázun	tin	a γápi	tu	patéra	tis	marías
admir:3PL the love:ACC the father:GEN the Mary:GEN							
'They admire the love of the father of Mary'							
b.	θavmdzun	tu	patéra	tin	a γápi	tis	marías
c.	tu	patéra	θavmázun	tin	a γápi	tis	marías
							NP-initial
							clause-initial

Heavy NP constituents, such as adpositional modifiers and relative clauses, may appear in clause-final position:

Modern Greek (Joseph and Philippaki-Warburton 1987: 113–14)

(38) a.	sístisa	mja	jinéka	pu	ti	γnórisa	s
introduced:1SG one woman:ACC COMP her:ACC met:1SG in							
to parísi s to jáni x γés							
the Paris:ACC to the John:ACC yesterday							
'I introduced a woman that I met in Paris to John yesterday'							
b.	sístisa	mja	jinéka	sto	jáni	x γés	pu ti γnórisa s to parísi

Extraposed relative clauses are also found in e.g. Dutch, Imbabura Quechua (Cole 1982: 51), and Sarcee (Cook 1984: 78).²⁶

Dutch

(39) Ik	heb	de	man	gezien	die	gisteren	mijn	fiets	gestolen	heeft
I have the man seen who yesterday my bike stolen has										
'I saw the man who stole my bike yesterday'										

The few examples of discontinuous constituents presented in this chapter should not be regarded as evidence that (true) discontinuity is an extremely rare phenomenon. It probably reflects the linguists' preoccupation with written language, whereas discontinuity is typically connected with the spoken variant. This is explicitly stated to be the case in e.g. Russian, in which it is a characteristic feature of everyday language (Gasparov 1978; Lönnqvist 1982: 349–50; Leinonen 1985: 125).

8.6. Conclusion

In this section I have presented some typological facts about the languages in the sample that are relevant for the discussion about word order in the final chapters of this book and I have explained how a number of (often interacting) general ordering principles (or 'competing motivations') account for the fact that only (p.260) a subset of all the logically possible ordering patterns are actually attested in the languages of the world. I have discussed some of the principles that have been formulated in the context of Dik's *Theory of Functional Grammar* (1997) and put forward a three-way typology of general ordering principles that is based on the *reference point* relative to which the position of a particular constituent is defined. Subsequently I dealt with three closely related issues: iconicity, domain integrity, and discontinuity. The *Principle of Domain Integrity* is one of the many ways in which a general iconic principle ('what belongs together semantically is expressed together in the actual linguistic expression') is reflected in morpho-syntactic patterns. It explains why on the whole discontinuity of constituents that are part of the same syntactic domain (here: NP) is a marked phenomenon (but more in the written than in the spoken mode), in terms of both frequency and phonological and morpho-syntactic coding devices (e.g. stress, case marking on the discontinuous constituent).

Both the *Principle of Domain Integrity* and the ordering principles to be discussed in Chapters 9 and 10 also play a role in the sentence, but since the present study is largely confined to simple, integral NPs this will not be investigated here.

An important advantage of the fact that we confine ourselves to simple, integral NPs is that it allows us to get a clearer picture of the functioning of the ordering principles that are discussed in this book: there is less competition and interaction between ordering principles in the simple NP than in complex NPs and clauses. For instance, in most languages the more pragmatically motivated principles are hardly relevant in the NP, whereas in the clause they play a rather prominent role (Tomlin 1986; Dik 1997). The following two chapters are devoted to the two other ordering principles that play an important role in the linear organization of constituents in the NP: the Principle of Head Proximity (a Type III principle) and the Principle of Scope (also a Type III principle).

Lastly, in Chapters 9 and 10 I will mainly concentrate on the languages in which demonstratives, numerals, and adjectives are all free, integrated modifiers of the noun phrase. Since most classifier languages lack a class of adjectives (section 4.3.5), this means that ordering patterns that include a classifier, as in (32) above, will be largely ignored. Such word order patterns are, however, discussed in Jones (1970), Goral (1978), and Rijkhoff (1990a); see especially (Greenberg 1975: 41) on the 'heavy statistical predominance' of the order *num clf* as against *clf num*.

Notes:

- (1) The idea of competing motivations has been taken up by several linguists: e.g. Greenberg (1966a); Steele (1975); Keenan (1979); Hawkins (1983, 1988a); DeLancey (1985); Du Bois (1985); Wierzbicka (1985); Tomlin (1986). See also Archangeli and Langendoen (1997) on constraint interaction in Optimality Theory and Croft (1990a: 56 f., 165, 192–7, 256) on the significance of competing motivations in typological work.
- (2) See also Gilligan (1987), who found that only seven languages in his 100-language sample did not have null-subjects in finite clauses.
- (3) See Gundel et al. (1988); Siewierska (1988: 8–14 1997a); Barker (1994: 34 f.); Croft (1995); Song (1991); Whaley (1997: chapter 6); and Dryer (1998) on the notions ‘basicness’, ‘dominance’, and ‘markedness’ in linguistic typology.
- (4) The element *m-o* is glossed as ActorFocus-Preterite; yet the translation has the verb in the present tense.
- (5) See Schroeder (1999: 189–99) for a detailed account of the ‘post-predicate position’ in Turkish, another non-rigid V-final language.
- (6) For a more detailed analysis of the Korean verb complex, see Kim (1987: 890–3) or Lee (1989: chapter 4).
- (7) Taking into account, of course, Greenberg’s Universal 6 (1966a: 79): ‘All languages with dominant order VSO have SVO as an alternative or as the only alternative basic order.’
- (8) Mithun (1992: 58) writes, for example, that in Cayuga (but also in other polysynthetic languages such as Ngandi and Coos) ‘the pronouns bear the primary case relations to the verb. The associated noun phrases function grammatically more as appositives to the pronominal affixes, rather than directly as verbal arguments themselves’ (on apposition, see also section 1.5.4.2); see McGregor (1997a: 54–87) for discussion on grammatical relations.
- (9) See also, for example, the discussion in Lewis (1967: 243 f.) on the problems concerning the V-final character of Turkish.
- (10) Relators (adpositions, case affixes, coordinators, and subordinators) are grammatical elements ‘which serve to link two constituents together, and/or to mark the function(s) of a constituent as specified in underlying clause structure. Relators may mark a relation of coordination or a relation of dependency’ (Dik 1997: 398, 406). Here I am only concerned with adpositions (pre- and postpositions) and case affixes. Mackenzie (1992a, 1992b) argues that e.g. English has a certain (closed) class of adpositions which are better regarded as lexical elements (predicates). On the case affixes, adpositions, and their word order typology, see e.g. Kahr (1976); Hawkins (1983); Limburg (1983); Kahrel (1985); Dotter and Strieder (1988); Tsunoda et al. (1995).
- (11) Most postpositions in Berbice Dutch Creole also occur as nominal predicates designating e.g. body-parts (Kouwenberg 1991: 143 f.; on body-part syntax see also Fox 1981).

(12) But cf. also Harriehausen (1990: 64–9) on relators in Hmong Njua.

(13) See Dik (1997: 154 f.), Siewierska (1991: 186–95); cf. also Blake (1994: 13–15, 51 f.).

(14) On agreement, as well as the coding of grammatical arguments and alignment systems in general, see Moravcsik (1978b); Nichols (1992); Siewierska (1996a,b, 1998).

(15) A possible exception is the *Principle of Iconic Ordering*, which refers to semantic relations within and across domains in general: ‘Constituents conform to [this principle] when their ordering in one way or another reflects the semantic content of the expression in which they occur’ (Dik 1997: 399).

(16) But cf. Dik (1997: 397) and Connolly (1983), who use templates to allocate constituents to particular positions.

(17) Under a different analysis *kinderen* could be an NP and *vrij veel* a quantifying satellite at clause level.

(18) In theory there are more options available than those given here, but it would take too long to explain why they are not feasible either.

(19) See also Greenberg (1966a: 102) on proximity hierarchies: ‘These hierarchies are presumably related to degrees of logical and psychological remoteness from the center, but no analysis is attempted here.’ On iconicity of distance, cf. also Croft (1990a: 174–83); Givón (1991: 86); and van Langendonck (1995: 84 f.). For some recent contributions and general discussion on iconicity in general, see e.g. Haiman (1980, 1985a, 1985b); Givón (1990: 966 ff.); Croft (1990a: 164–92); Landsberg (1995).

(20) In languages where the attributive demonstrative pronoun requires the presence of an article, these two constituents usually occur in adjacent positions in the NP, with the article in between the demonstrative and the numeral (cf. Van Valin and LaPolla 1997: 62). In such cases I would argue that one of two things is the case (allowing of course for the fact that a language may be changing its grammatical system and that grammaticalization is a gradual process). Either the demonstrative is not an integral part of the NP (cf. Moravcsik 1997: 319) or the article is no longer (only) a marker of definiteness and has changed into a Stage II article (section 3.3.2.4; Greenberg 1978b).

(21) Greenberg (1975) argues that, diachronically, the appositional numeral classifier precedes the integrated prenominal variant, ‘at least in those languages in which the classifier construction arises through internal processes’. He proposes a process whereby the numeral classifier phrase starts out as a clausal constituent (the ‘adverbial’ construction), which then becomes an appositional phrase, and which may ultimately end up as a fully integrated constituent in prenominal position.

(22) Recall that in many numeral classifier languages the numeral classifier phrase is an appositional construction, meaning that it is *not* part of an *integral* NP (section 1.5.4).

(23) Cf. McGregor (1989: 219) on the difference between fractured and discontinuous NPs in Gooniyandi.

(24) Australian languages outside the sample whose NPs have been analyzed as appositional structures are, for example, Mangarayi (Merlan 1982: 49), Ngiyambaa (Donaldson 1980: 232), and Walpiri (Simpson 1983; Hale 1983).

(25) Recall that quantifiers are excluded in this study. In many languages the universal quantifier ‘all’ may occur both inside and outside the NP; this phenomenon is usually referred to as ‘quantifier float’ (e.g. Bouma 1987). But here, too, it is often questionable whether the universal quantifier is part of an integral NP.

(26) Cf. also Ziv and Cole (1974); Mallinson (1986).

9 Greenbergian Word Order Correlations and the Principle of Head Proximity

9.1. Introduction

The *Principle of Head Proximity* (Rijkhoff 1984, 1986, 1990a) is an interpretation of the word order facts as presented in Greenberg's milestone article on language universals (Greenberg 1966a) and in Hawkins's *Word Order Universals* (1983), which is a contribution to Greenberg's data. Table 9.1 summarizes the data in Hawkins's *Expanded Sample* (Hawkins 1983: 281–90). As in Hawkins's original classification languages are classified as V-1 (V-initial), SVO, or SOV (A = Adjective, G = Genitive/Possessor NP, and ‘--’ means both A and G appear before or after the noun).

These data demonstrate that there is a general tendency across languages to avoid having adjectives (A) and possessor NPs (G) between the head of the NP (N) and the head of the clause (V), and that this tendency is stronger for the possessor NP (G, an embedded modifier) than for the adjective (A, a non-embedded modifier). The data also show that in the V-1 and SOV languages a possessor NP (G) is only permitted in between heads V and N if an adjective can appear there as well.

Table 9.1. Classification of languages

(1)	V-1 & N--	38	(5)	SVO & N--	56	(9)	SOV & N--	21
(2)	V-I & ANG	13	(6)	SVO & ANG	17	(10)	SOV & ANG	0
(3)	V-I & --N	2	(7)	SVO & --N	19	(11)	SOV & --N	98
(4)	V-I & GNA	0	(8)	SVO & GNA	17	(12)	SOV & GNA	55
		53			109			174

Source: Hawkins (1983: 288).

Thus, most V-1 languages (type 1) have neither the adjective nor the possessor NP in between heads V and N, and in the second largest group (type 2) only the adjective precedes the noun. Hawkins's sample has no verb-initial languages (p.262) with G preceding and A following the head and only two languages in which both A and G occur in postnominal position.

V-initial languages

Type (1) V[N--]NP... (38 1gs.)

Type (2) V[ANG]NP... (13 1gs.)

Type (3) V[-N]NP... (2 1gs.)

Type (4) V[GNA]NP... (0 1gs.)

Similarly, the largest class of SOV-languages is the group in which both A and G do not appear between the head noun and the main verb V (type 11), and the second largest group has only the adjective in between heads N and V (type 12). In a small group both A and G follow the noun (type 9), but in no verb-final language does A precede and G follow the noun (type 10).

V-final languages

Type (11): ... [-N]NP V (98 lgs.)

Type (12): ... [GNA]NP V (55 lgs.)

Type (9): ... [N--]NP V (21 lgs.)

Type (10): ... [ANG]NP V (0 lgs.)

SVO-languages display properties of both V-initial and V-final languages in that the second largest group (34 lgs.) permits either GNA or ANG order in the NP (types 6 and 8); hence Dik (1997: 411) speaks of Postfield (V-initial) and Prefield (V-final) derived SVO-languages. Nevertheless the largest group of SVO-languages (type 5: 56 lgs.) resembles the best attested verb-initial type in that they have both A and G following the noun. At first sight it looks as if languages of type 5 do not follow the general tendency to avoid having the two modifiers A and G between the head of the clause V and the head of the noun phrase N. Since in an SVO-language the verb is normally preceded by the subject, modifiers in the subject NP would occur between heads N and V, as in: [N--]NP V ... (type 5). In practice, however, subject NPs with adnominal modifiers are the exception rather than the rule, because cross-linguistically subjects (and transitive subjects in particular) are associated with givenness and humanness (Givón 1981b). This means that the subject position is most typically not occupied by a lexical NP but rather by a pronoun or a proper name (for some empirical evidence see Rijkhoff 1986: 121–2). Recall also that many languages are so-called pro-drop languages, meaning that the subject is often only marked by a morphologically bound pronominal element in the verbal complex, in which case the preverbal subject position in the clausal structure is simply left unused (this is indicated by the round brackets in the listing of the SVO types below; cf. also section 8.2.1).

(p.263)

V-second languages

Type (5): (NPs)V[N--]Np ... (56 1gs.)

Types (6, 8): (NPs)V[-N-]NP... (34 1gs.), [-N-]NP = [ANG]NP or [GNA]NP

Type (7): (NPs)V[--N]NP... (191gs.)

The observed word order tendencies in Hawkins's *Extended Sample* are captured in the following ordering principle:

- (1) In the noun phrase there is a tendency to avoid the placement of adnominal adjectives (A) and possessor NPs (G) between the head of the noun phrase (N)

and the head of the clause (V); this tendency is stronger for the possessor NP (an embedded modifier) than for the adjective (a non-embedded modifier).

Keeping in mind the difference between embedded and non-embedded modifiers (see also sections 1.5.4.3 and 7.4), one could hypothesize that what holds for possessor NPs and adjectives also goes for other modifier categories such as determiners, numerals, and relative clauses:

(2) In the noun phrase there is a tendency to avoid the occurrence of adnominal modifiers between the head of the noun phrase (N) and the head of the clause (V); this tendency is stronger for embedded modifiers than for non-embedded modifiers.

The data in Table 9.1 can also be interpreted in positive terms and be regarded as evidence that across languages there is a tendency to reduce the distance between the head of the noun phrase and the head of the clause. This is captured in the following version, which I call the *Principle of Head Proximity I* (i.e. the weak version).¹

(3) *The Principle of Head Proximity I*: In the noun phrase the preferred position of the head noun N is as close as possible to the head of the clause V. This cross-linguistic tendency is reflected in the position of adnominal modifiers, which tend not to occur between the head of the noun phrase N and the head of the clause V; this holds in particular for embedded modifiers such as possessor NPs and relative clauses.

(p.264) This weak version of the *Principle of Head Proximity* only refers to heads N and V, but since adjectives and adverbs are also head constituents (i.e. lexical elements or predicates), one could go one step further and formulate a strong version of the *Principle of Head Proximity* (version II), which generalizes across all types of head constituents (verbs, nouns, adjectives, and adverbs):

(4) *The Principle of Head Proximity II*: In a subordinate domain, the preferred position of the head constituent is as close as possible to the head of the superordinate domain.

The strong version assumes that domains (clauses, noun phrases, adjective phrases, adverb phrases) are in the following, hierarchical relationship with each other (I will restrict myself to degree adverbs here):²

(5) SUPERORDINATE SUBORDINATE DOMAIN(S) DOMAIN

clause (head: V)	noun phrase (head: N), embedded clause (head: V)
noun phrase (head: N)	adjective phrase (head: A), embedded adnominal modifier such as a possessor NP (head: N) or relative clause (head: V)
adjective phrase (head: A)	degree adverb (head: adv)

On the basis of the second (strong) version of the *Principle of Head Proximity* we can now formulate two main hypotheses regarding the syntax of adnominal modifiers (on *head proximity* and syntax at the level of the clause, see section 9.3.3.3 below and Rijkhoff 1986, 1984; see also Appendix to Chapter 9).³

First, the strong version of the *Principle of Head Proximity* predicts that the preferred position of the head of any lexical modifier is immediately before or after the noun, but

that short lexical modifiers (such as adjectives) are preferred closer to the noun than long, embedded modifiers (such as possessor NP or relative clause) if they both occur on the same side of the noun, since this yields better Head Proximity (on the relative order of adjective and embedded modifier, see section 9.3.4.2.2).

(6)

- a. N A [HEAD...]_{G/Rel} (preferred)
 
- b. N [HEAD...]_{G/Rel} A (not preferred)
 

(p.265)

-
- (7) a. [...HEAD]_{G/Rel} A N (preferred)
-
- b. A [... HEAD]_{G/Rel} N (not preferred)

According to the *Principle of Domain Integrity*, too, the preferred position of *embedded modifiers* is not in between constituents of the matrix NP, i.e. in the periphery of the matrix domain. In other words, when a noun is modified by an adjective and an embedded modifier (and they occur on the same side of the noun), the combination of the *Principle of Domain Integrity* and the *Principle of Head Proximity* only ‘permits’ the head of the (short) non-embedded modifier, the adjective, to be adjacent to the noun. Of course, in the simple NP (the central topic of this book), where the adjective is the only lexical adnominal modifier, we also expect the adjective to occur immediately before or after the noun. This hypothesis will be tested in section 9.2.

Secondly (this has already been mentioned above) the strong version of the *Principle of Head Proximity* implies that the preferred position of *any* adnominal modifier is not between head noun N and the head of the clause V, still keeping in mind that the tendency to avoid having adnominal modifiers between heads N and V is stronger for embedded than for non-embedded modifiers. Presumably this is so because embedded modifiers tend to be longer than non-embedded modifiers and may contain other embedded modifiers; consequently to have an embedded modifier between N and V would be worse in terms of ‘head proximity’ than a non-embedded modifier. This hypothesis will be investigated in some detail in section 9.3.

Before we continue, it may be useful to explain in what sense the *Principle of Head Proximity* is an elaboration of Behaghel’s first law, which was mentioned in Chapter 8 and which can be paraphrased as: ‘what belongs together semantically, occurs together syntactically.’ The reason why I regard the *Principle of Head Proximity* as a more specific manifestation of Behaghel’s general iconic principle is that it suggests that there is a close semantic bond between heads of hierarchically ordered domains. Semantically the head constituent is the most informative constituent of a domain and the data presented above indicate that linguistic structures are easier to process when it does not take too much time and effort to identify the head of the next subordinate or superordinate domain. This close bond between heads of hierarchically ordered domains, then, is iconically reflected in the cross-linguistic tendency not to

have (too many) modifying elements in between these heads. Evidence for the central role that heads play in language processing strategies can be found, for example, in Clark and Clark (1977: 53; note that what they call ‘constituent’ is called ‘domain’ in my terminology):

Listeners have at their command a battery of mental strategies by which they segment sentences into constituents, classify them, and construct semantic representations from them. These strategies rely on the fact that sentences contain elements listeners can use as clues to proper segmentation.

(p.266) One such strategy, according to Clark and Clark (1977: 61), goes as follows (cf. also Moore 1972):

Strategy 2: After identifying the beginning of a constituent look for content words appropriate to that type of constituent.

In the light of the evidence presented above, one could elaborate on Clark and Clark’s second strategy in the following way: whenever you identify a domain, look for the head constituent of that domain, and then try to relate this constituent to the head of the superordinate domain (in a head-final language) or the head of the subordinate domain (in a head-initial language). The idea that linguistic expressions are easier to process when the distance between heads of domains is short rather than long has also been defended in Frazier (1985: 146) when she investigated to what extent certain basic word orders would complicate sentence processing:⁴ ‘If the head of a phrase and the head of its complement were separated, then the phrasal packages constructed by the limited-window procedure would not form a coherent semantic unit.’ The two hypotheses put forward on the basis of the *Principle of Head Proximity* will be tested in the next two sections (9.2 and 9.3).

9.2. The first hypothesis: noun-adjective adjacency

This section focuses on the first hypothesis that was formulated on the basis of the *Principle of Head Proximity*, namely that the adjective always appears next to the head noun. To check this hypothesis, I only took into account the twenty-five languages in the sample which were judged to have a category of adjectives (i.e languages of types 1, 2, 3, and 3/4); these languages were enumerated in Table 4.3.

9.2.1. Non-adjacent adjectives in languages with a major class of adjectives

Evidence that the adjective and the head noun are not (always) contiguous was found in descriptions of four languages that have a major (distinct or flexible) class of adjectives: Bukiyp, Hungarian, Turkish, and Nasioi.

9.2.1.1. Bukiyp

It is not certain that Bukiyp is actually a counter-example to the first hypothesis based on the *Principle of Head Proximity*. According to the general NP schema provided by Conrad (1991: 57) a possessive modifier (NP or pronoun) occurs between the (prenominal) adjective and the noun. However, when the noun is modified by both an adjective and a possessive modifier, the latter is ‘permuted to postnominal position’ (*ibid.*). The same schema indicates that a derived adjective(p.267) must occur after

the head noun in Bukiyp (as in the example below), and that this position can also be occupied by a basic adjective.

Bukiyp (Conrad 1991: 59)

(8) apak-i-ny	bolany	généga-i-ny
our-POS-CL8.SG	talk	indigene-POS-CL8.SG
'our indigenous language'		

It is not clear whether a postnominal adjective (whether basic or derived) can occur with a postnominal possessive modifier, and, if so, how they are ordered relative to the head noun. The rather extensive list of examples following the schema does not contain such a combination of modifiers.

9.2.1.2. Hungarian

The adjective and the noun are normally contiguous in Hungarian, but in an exclamation (i.e. not in NPs in a regular sentence) the emphasized adjective may precede an indefinite article or a numeral:

Hungarian (Tompa 1968: 77)

(9) gyönyörű	egy	állat!
beautiful	an	animal
'a beautiful animal!'		

Hungarian (Lotz 1939: 231)

(10) szép	két	ló!
beautiful	two	horse
'two beautiful horses!'		

This ordering may be explained if we assume that the first position in an NP is reserved for constituents with a special (pragmatic) function, just like the clause-initial position. This is by no means unusual, as can also be shown by examples from e.g. Babungo in which an emphasized demonstrative or possessive pronoun may occur in NP-initial position, whereas they normally appear after the noun (Schaub 1985: 77).

9.2.1.3. Turkish

At first sight it seems a similar situation obtains in Turkish (and Tamil, see below).⁵ In Turkish the adjective may precede or follow *bir* 'a, one', but if *bir* serves as a cardinal numeral ('one') it must precede the adjective. According to Krámský (1972: 111–12), referring to Dmitrijev (1939: 49), *bir* only occurs in between the adjective and the head noun when the adjective is emphasized:⁶

(p.268) Turkish (Krámský 1972: 111–12)

(11)	bir	büyük	ev
	a/one	big	house
'a/one big house'			

(12)	büyük	bir	ev
	big	a	house
'a big house'			

Should one wish to stress both the indefiniteness and the adjectival property of the referent, then *bir* appears twice:⁷

Turkish (Krámksy 1972: 112)

(13)	bir	güzel	bir	kız
	BIR	beautiful	BIR	girl
'a beautiful girl'				

Other studies, however, indicate that indefinite *bir* can also be analyzed as a marker of individuality (i.e. as a kind of nominal aspect marker) rather than as an indefinite article; therefore I will return to the case of Turkish indefinite *bir* in section 10.2.1.2.4.

9.2.1.4. Nasioi

Nasioi demonstratives may occur in various places in the NP, including the position between the noun and the adjective (Rausch 1912: 132; CI5= noun class 15).

Nasioi (Rausch 1912: 132)

(14)	koinave	ave	itikave
	tree:C15	DEM:C15	high:C15
'this high tree' (or: 'this tree is high'; see note 8)			

It was already mentioned in section 4.3.1 that due to the elaborate noun class system adjectives can also occur alone and the same holds for demonstratives and numerals. Class suffixes are generally identical to the class marked forms of the (p.269) numeral 'one', and they also seem to serve as nominalizing or anaphoric elements.⁸ The following examples show that demonstratives, numerals, and adjectives may occur without the head noun if coded for class: in such cases the presence of a class marker alone suffices to communicate what (kind of) object is being referred to. Compare (a class 15 (CI5) suffix is used in connection with wooden objects):

Nasioi (Rausch 1912: 122)

(15)	koinave	peténave
------	---------	----------

tree:C15 soft:C15

'the soft wood/tree'

When the class suffix appears on the adjective, it need not occur on the noun as well; in fact, in such cases the noun can be left out altogether (as in (17)).

(16) koi peténave

tree soft:C15

'the soft wood/tree'

(17) peténave

soft:C15

'the soft wood/tree'

Compare also the following examples with a demonstrative:

Nasioi (Rausch 1912: 131–2)⁹

(18) koinave ave orave

tree:C15 DEM:C15 bad:C15

'this bad tree'

(19) koinave orave ave

tree:C15 bad:C15 DEM:C15

'this bad tree'

(20) ave koinave orave

DEM:C15 tree:C15 bad:C15

'this bad tree'

(21) ave orave

DEM:C15 bad:C15

'this bad one (=tree)'

(p.270)

(22) orave ave

bad:C15 DEM:C15

'this bad one (=tree)'

(23) ave

DEM:C15

'this one (=tree)'

In sum, although much remains unclear about the structure of the Nasioi NP, these examples seem to indicate that both adjectives and demonstratives (when provided with the appropriate class marker) are independently referring expressions and (syntactically speaking) not dependants of the noun. This means that, in combination with a noun, demonstratives and adjectives (plus class affix) can be regarded as instances of appositive modification. In fact, such an analysis would be supported by the great variety of ordering possibilities, as shown in the examples above.¹⁰

9.2.2. Non-adjacent adjectives in languages with a minor class of adjectives

Each of the three languages to be discussed below (Babungo, Oromo, Tamil) only has a very small number of basic (underived) adjectives. Babungo has eleven of them; the precise number of adjectives in the other two languages is not given, but Oromo is said to have 'a shortage of adjectives' (Bender et al. 1976b: 145), and Asher (1982: 187) states that Tamil has 'a very small handful' of them. Unfortunately it is not indicated for any of these three languages whether all (i.e. basic and non-basic) adjectival modifiers need not immediately precede or follow the head noun, or whether the various forms and constructions that are collectively referred to as 'adjectives' have different syntactic properties. In other words, it may be the case that in some languages cases of non-adjacency are restricted to modifiers that I do not regard as true (basic) adjectives, but rather as embedded modifiers (whose syntactic properties differ from those of non-embedded modifiers, see the *Principle of Domain Integrity*).

(p.271) 9.2.2.1. Babungo

As a rule the Babungo adjective immediately follows the head noun, except in the case of inalienably possessed nouns, such as those denoting body-parts and kinship relations. Compare:

Babungo (Schaub 1985: 77)

(24) ɳkáw jéə kāŋ

chair good my

'my good chair'

(25) zú wī jéə

wife his good

'his good wife'

The best explanation for this (apparent) exception to the *Principle of Head Proximity* is in fact provided by the author himself (Schaub 1985: 77), who suggests that the noun

and the possessive pronoun can be conceived of as forming a single constituent ('a closely knit unit functioning as a whole as the head'), thus reflecting the strong conceptual bond between the inalienably possessed item and its possessor in the actual linguistic expression. It is not clear from the grammar to what extent the possessive forms in question can be treated as true suffixes, but we will see below (section 9.4.2.2) that if we are indeed dealing with suffixed forms, Babungo would fit the cross-linguistic norm in that the bound variant is used to express inalienable possession in those languages which have both free and bound possessor pronouns. In any case, if in Babungo the noun + pronoun together form the head constituent of the NP (as suggested in Schaub's grammar of Babungo), then this construction does not constitute a counter-example to the *Principle of Head Proximity*.

9.2.2.2. Oromo

In Oromo, according to Bender et al. (1976b: 144), adjectives, numerals, and demonstratives all follow the noun in that order, but Stroomer (1987: 59), whose book is confined to the three Oromo dialects Boraana, Orma, and Waata, states that the numeral may also appear before the adjective:

Orma Oromo (Stroomer 1987: 59)

(26) gaala lamaani gugurdaami kuni sookoo ḫak-e
camel two big:S these:S market go-3SG.M.PAST
'These two big camels went to the market'

Interestingly Bender et al. (1976b: 144) added that, for example, *namni haman* [man:S bad:S] 'the bad man' and *nitin hamtun* [woman:S bad:S] 'the bad woman' 'may be thought of as phrases with the second word as a comment on the first, i.e. a kind of apposition: "the man, the bad (one)"; "the woman, the bad (one)".' This means that a more literal translation of the subject NP in the example from (p.272) Orma Oromo could read (recall that Oromo has no definite article): '(the) two camels, these big (ones)' or '(the) camels, these two big (ones)' or '(the) camels, (the) two (ones), these big (ones)' or even '(the) camels, (the) two (ones), (the) big (ones), these (ones)'.¹¹

The only other (free) modifier that may appear between the noun and the adjective is the possessor NP (cf. sections 4.3.2 and 9.4.2 below on doubling of A and G in Oromo), but in his description of the Wellegga dialect of Oromo Gragg (1976: 191) notes that 'if the possessive is marked with *kan* it comes after the adjective' (the additional marker *kan* in (28) is formally identical to the relative pronoun; see Bender et al. 1976b: 137 and Gragg 1976: 191). Compare:

Wellegga Oromo (Gragg 1976: 191)

(27) mana mootii guddaa
house king:GEN big
'the big house of the king' or 'the house of the big king'

(28) mana guddaa kan mootii

house big KAN king:GEN

'the big house of the king'

Thus, with the adjective following the possessor NP, as in (27), 'big' may be interpreted as a modifier of 'house' or 'king', but with the order reversed, as in (28), 'big' can only modify 'house'. Although it is not possible to give a fully satisfying explanation on the basis of the information that is available, these facts do seem to suggest that the non-adjacent adjective without *kan* is indeed an apposed rather than an integral modifier of the noun (see above). In the literal paraphrase of the non-adjacent variant in (27), i.e. 'the king's house, the big (one)', the adjective is neither an integral part of the NP headed by 'house' nor an integral part of the possessor NP headed by 'king'; hence it can relate to both 'house' and 'king'.

9.2.2.3. Tamil

In Tamil adjectives normally occur next to the head noun, but if emphasized they may be placed before an article, demonstrative, or quantifier (Asher 1982: 66, 94):

Tamil (Asher 1982: 66)

(29) ażakaana oru cinna viṭu

beautiful one small house

'a beautiful little house'

(30) ażakaana anta putu poṭave

beautiful that new sari

'that beautiful new sari'

(p.273)

(31) rompa oyaramaana naalu tennamaram

very tall

four coconut_tree

'four very tall coconut trees'

Apparently in Tamil, too, emphasized adjectives may appear in NP-initial position, just like in e.g. Hungarian (see above).

9.2.3. Non-adjacent adjectives, Universal 20, and asymmetry

The hypothesis that a true attributive adjective is normally adjacent to the head noun is generally confirmed by the data in the sample languages. We saw that the sample contains about half a dozen languages in which another modifier can appear between the adjective and the noun, but as far as the evidence goes all cases of non-adjacency could be attributed to special pragmatic conditions (i.e. emphasis, as in Hungarian,

Turkish, Tamil), to apposition (possibly Nasioi), or to what might be called ‘compound headedness’ in that, according to Schaub (1985: 77), the noun + possessor_pronoun construction in Babungo can be regarded as a single unit (recall that the case of Bukiyip is probably not relevant). The only problematic case is provided by Oromo, for which insufficient data are available to decide whether or not this language is a true counter-example to the first hypothesis based on the *Principle of Head Proximity*.

Interestingly the Oromo dialects are all spoken in a region in Africa (roughly covering the central and eastern part of the continent) where quite a few other languages can be found in which adjectives do not necessarily appear immediately before or after the noun. In fact, when Greenberg (1966a: 86–7) presents his *Universal 20*, which reads:

Universal 20. When any or all of the items (demonstrative, numeral, and descriptive adjective) precede the noun, they are always found in that order. If they follow, the order is either the same or its exact opposite,

he mentions a language from the same area, Kikuyu (a Central Bantu language spoken in Kenya), as the only language in his sample with the order [N dem num A].

Some twenty years later Hawkins (1983: 119–20) reformulated Greenberg’s *Universal 20*, after it had come to his attention that the orders [N A num dem] and [N dem num A] are not the only possibilities when demonstrative, numeral, and adjective all follow the noun. Referring to Hyman (1979, 1981), he mentions two more Bantu languages with a deviant pattern: Aghem [N A dem num] and Noni (which besides [N dem num A] also has [N dem A num]), both spoken in Cameroon.

Universal 20'. When any or all of the items (demonstrative, numeral, and descriptive adjective) precede the noun, they (i.e., those that do precede) are always found in that order. For those that follow, no predictions are made, though the most frequent order is (p.274) the mirror-image of the order for preceding modifiers. In no case does the adjective precede the head when the demonstrative or numeral follow.

Three years before the publication of Hawkins’s book, Heine (1980) had also listed a number of languages (all spoken in Kenya) with ordering sequences that deviate from the ones deemed possible according to Greenberg’s *Universal 20*. The word order sequences given below are not the only possible patterns; all the languages mentioned are said to display an enormous amount of variation.

Rendille (Afro-Asiatic; Cushitic):	N-g-dem num A G Rel
Gabra (Afro-Asiatic; Cushitic):	N g num A dem G Rel
Sampur (Nilo-Saharan; Eastern Nilotic):	dem N g num A G Rel
Camus (Nilo-Saharan; Eastern Nilotic):	dem N g num A G Rel
Turkana (Nilo-Saharan; Eastern Nilotic):	N dem g num A G Rel
Luo (Nilo-Saharan; Western Nilotic):	N-g num A dem G Rel
Logoli (Niger-Congo; Central Bantu):	N g num A G dem Rel

To this list we may add Nkore-Kiga (Taylor 1985: 55), another Central Bantu language (cf. Rijkhoff 1990a: 17–19, 30–1):

Nkore-Kiga: N g dem A/appositives num A_{verbal} Rel num

Although detailed analyses of NPs in the languages listed above are not available (Heine's article is only six pages long), it is still possible to detect some common patterns, which may explain why the noun and the adjective need not be adjacent. First of all we see that, as in Babungo, certain modifiers in the languages that Heine listed do not occur as free forms, but only as nominal suffixes (cf. Rendille and Luo). Secondly, many of the languages in the area (including at least some of the languages mentioned above) lack a distinct class of adjectives; instead adjectival notions are expressed by nouns or verbs (see the remarks by Bender et al. 1976b: 145 and Sasse 1976: 206 in section 4.3.2). For example, adjectives in e.g. Sampur and probably also in Camus (both are dialects of Nilo-Saharan Maa) are basically verbs heading a relative construction (Heine 1980: 182), and at several points in his description of Nkore-Kiga Taylor repeats statements to the effect that 'the true adjective has a form and function similar to that of a noun' (Taylor 1985: 85; see also p. 174, p. 195).¹² Note furthermore that in Babungo and other Bantu languages numerals are often nominals (section 5.3.2).¹³ In all these cases we are dealing with embedded noun modifiers, which according to the *Principle of Domain Integrity* should not occur in between constituents of the matrix NP (sections 8.4.2. and 8.5). Thirdly, it may be the case that certain (p.275) modifiers are not part of the integral NP but rather appositions; in fact this has also been proposed by Seiler (1978a: 322–3) in his discussion of the Kikuyu NP (see also Bender et al. 1916b 144 on the appositional nature of Oromo adjectives in section 9.2.2.2). Note that in Beja (an isolate in the Cushitic family and supposedly the most archaic language of the Cushitic branch; cf. Zaborski 1976: 70) all modifiers seem to be able to occur in an appositional construction (Hudson 1976: 106):

modifiers may either precede or follow the noun ... The modifiers show concord of number, gender, and case with the noun in either position, and if they follow they also show concord of definiteness, as in Arabic [recall that Oromo does not mark (in)definiteness—JR]:/ták wìn/a big (wìn) man (ták)' but /?uu-tak ?uu-wìn/'the big man'.¹⁴

Appositional forms of modification would also at least partly explain the enormous degree of syntactic variation in the languages mentioned in Heine's article (1980). Note finally that the order of postnominal constituents in the languages listed above seems to be influenced by the *Principle of Increasing Complexity* (or LIPOC), since we see a strong tendency to place 'heavier' constituents towards the end of the NP (Dik 1997: 404, 411; cf. section 8.4.1).

In sum, although we know very little about head-modifier relationships in the NPs of the languages referred to above, there are several indications that we are either dealing with non-simplex NPs (embedded modifiers, as in for example Sampur, Camus, Dasenech) or non-integrated modifiers (appositional modification; possibly Kikuyu, Oromo) when adjectives are not adjacent with the noun. The appositional analysis is supported by the fact that all cases of noun-adjective discontinuity presented here involve orderings in which the adjective and nearly all other modifier categories appear after the noun. Of course, this asymmetry between pre- and postnominal modification in the NP was precisely the reason why Greenberg and later Hawkins could give a single

ordering pattern for prenominal modifiers, but had to allow for two or more variants for the relative order of postnominal noun modifiers. Since appositional modifiers are not part of an integral NP structure one could hypothesize that LIPOC and other ordering forces counteracting the *Principle of Head Proximity* have a bigger impact on modifiers that are not part of an integral NP than on fully integrated noun modifiers; this would explain the tendency to place lighter constituents before heavy constituents (i.e. dem > num > A > G > Rel).

So far I have discussed languages in which adjectives do not normally appear next to the noun, which runs counter to the first hypothesis formulated on the basis of the *Principle of Head Proximity*. Apart from cases in which strong pragmatic forces are involved (as with exclamations; see section 9.2.1.2), all (p.276) counter-examples to the *Principle of Head Proximity* are restricted to languages in which the adjective follows the noun. There are various indications, however, that we may only be dealing with apparent counter-examples in these cases; for instance, because intervening modifiers are nominal suffixes (e.g. Babungo) or because adjectives are in fact verbs heading a relative clause (e.g. Sampur), whose ordering properties differ from those of adjectives. Furthermore, since appositional modifiers typically occur after the head noun, it was hypothesized that the asymmetry in the ordering possibilities of pre- and postnominal adjectives (and certain other noun modifiers—see above) might be due to the adjective not being a fully integrated noun modifier (see also section 9.5 on asymmetries concerning the syntax of embedded noun modifiers).¹⁵

9.3. The second hypothesis: the position of modifiers relative to the noun

The second hypothesis that was formulated on the basis of the *Principle of Head Proximity* pertains to the position of modifiers in the NP relative to the head noun. In languages where NPs normally follow the verb (V-initial and SVO-languages) the preferred position of adnominal modifiers would be after the noun; conversely, in languages where NPs normally precede the verb (SOV-languages) adnominal modifiers would tend to precede the noun. That is, adnominal modifiers will tend to avoid occurring between the head of the noun phrase N and the head of the sentence V. Recall that it was postulated on the basis of Hawkins's data (Table 9.1) that this tendency is stronger for embedded domains (possessor NP, relative clause) than for non-embedded domains (demonstrative, numeral, adjective), the reason being, presumably, that embedded modifiers (which can have their own embedded domains) generally take up more space than non-embedded domains, i.e. their presence between N and V would interfere with 'head proximity' more than the presence of non-embedded modifiers.

Since correlations between noun-modifier combinations and other ordering patterns (such as verb and object NP) have been investigated before, I will first discuss some previous work in this area.

9.3.1. Word order correlations

In 1961, at the Conference on Language Universals in Dobbs Ferry (New York), Greenberg presented word order data which suggested (among other things) that if one knew the basic word order of a language, one could always or

with (p.277) considerable reliability predict certain other ordering patterns in that language (Greenberg 1966a). For instance:

Universal 2. In languages with prepositions, the genitive almost always follows the governing noun, while in languages with postpositions it almost always precedes.

Universal 3. Languages with dominant VSO order are always prepositional.

Universal 4. With overwhelmingly more than chance frequency, languages with normal SOV order are postpositional.

Some two decades later Hawkins (1983: 65, 82), using a sample containing over 300 languages, elaborated on Greenberg's original work and formulated some new—often exceptionless—universals, such as (cf. Table 9.1):¹⁶

If a language has VSO order, then if the adjective follows the noun, the genitive follows the noun; i.e. $VSO \supset (NA \supset NG)$.

If a language has noun before genitive, then it has noun before relative clause; i.e., $NG \supset NRel$ (equivalently: $RelN \supset GN$).

However, recent large-scale research on Greenbergian word order correlations by Dryer (1991, 1992) produced several counter-examples to some of Hawkins's implications that were deemed to be exceptionless (see Table 9.2). To this list of counter-examples could be added Tigre (Prep & SOV & AN & NG & RelN) and Tsou (V-1 & NA & GN).¹⁷

Table 9.2. Exceptions to Hawkins's universals

Universal (Hawkins 1983)	Exception	Type
$V-I \supset (NA \supset NG)$	Kilivila	VOS & NA & GN
	Garawa	VOS & NA & GN
	Yagua	VSO & NA & GN
	Guajajara	VSO & NA & GN
$Prep \& -SVO \supset (NA \supset NG)$	Kilivila	Prep & VOS & NA & GN
	Garawa	Prep & VOS & NA & GN
$Prep \& -SVO \supset (NDem \supset NG)$	Kilivila	Prep & VOS & NDem & GN

Source: Dryer (1991: 450).

(p.278) 9.3.2. Dryer's correlation pairs

Dryer's sample mentioned in the previous section contained over 600 languages (but see n. 1 in Dryer 1992: 83) and it was his aim to test which word order pairs actually correlate with the order of object NP and verb (VO/OV). Claims concerning the existence of such correlation pairs had been made in a number of publications following Greenberg's original research (notably by Lehmann 1973, 1978a, and Vennemann 1973, 1976), but they had never been substantiated by systematic evidence. Dryer's investigation shows that there are indeed several word order pairs that more or

less correlate with the VO/OV distinction, among them the pair noun-possessor NP and the pair noun-relative clause (both embedded noun modifiers). As was hypothesized on the basis of the *Principle of Head Proximity* he found that ‘RelN order is more common among OV languages than it is among VO languages, and conversely for NRel order’ (Dryer 1992: 86). The same holds for the order of noun and possessor NP (G): VO-languages tend to have NG order and there is even a ‘strong preference’ for GN order in OV-languages. However, Dryer (1992: 95–6) did not find evidence for a correlation between VO/OV order and the order of noun and adjective, or the order of noun and demonstrative. Since the data about the order of noun and numeral were more difficult to interpret, he decided to leave this pair ‘unclassified’ (Dryer 1992: 118). His results showed that ‘the two orders of numeral and noun are equally common among OV languages’, but that ‘outside of Africa, VO languages exhibit a strong tendency to be NumN’ (*ibid.*). Thus, only in Africa do we find the preferred pattern in terms of head proximity, i.e. ‘a very strong tendency’ for the order NNum in VO-languages. However, Dryer’s discussion of the numeral + classifier construction in Burmese (which I regard as an apposition; cf. 5.2.2.1.3) shows that he did not take into consideration the fact that in Burmese and many other languages certain elements are not part of the (simple) integral NP. Thus it remains to be seen what his results would have been if he had distinguished between elements that are part of the integral NP and those that are not, and between grammatical and lexical instances of adnominal numerals, and between true adjectives and embedded structures (headed by an abstract noun or a stative verb). In fact I will argue in section 9.3.3.4 that some of Dryer’s categories are too heterogeneous to yield reliable results.¹⁸

Table 9.3 shows that the complete list of ordering pairs that more or less correlate with the VO/OV distinctions according to Dryer’s investigation include more than just a few noun-modifier combinations. To account for these correlation pairs, Dryer proposed his *Branching Direction Theory*(BDT), according to which there is a tendency for phrasal categories to precede non-phrasal (p.279)

Table 9.3. Dryer’s correlation pairs

Verb patterners	Object patterners	Example
verb	subject	(there) entered + a tall man
adposition	NP	on + the table
copula verb	predicate	is + a teacher
‘want’	VP	wants + to see Mary
tense/aspect aux. verb	VP	has + eaten dinner
negative auxiliary	VP	(see Dryer 1992: 101)
complementizer	s	that + John is sick
question particle	S	(see Dryer 1992: 102–3)
adverbial subordinator	s	because + Bob has left
article	N'	the + tall man
plural word	N'	(see Dryer 1992: 104–5)
noun	genitive	father + of John
noun	relative clause	movies + that we saw

adjective	standard of comparison	taller + than Bob
verb	PP	slept + on the floor
verb	manner adverb	ran + slowly

source: Dryer (1992: 108).

categories in OV-languages and vice versa in VO-languages (Dryer 1992: 132–3). The basic version of the BDT reads as follows (Dryer 1992: 87, 109):

(32) BRANCHING DIRECTION THEORY: Verb patters are non-phrasal (non-branching, lexical) categories and object patters phrasal (branching) categories. That is, a pair of elements X and Y will employ the order XY significantly more often among VO-languages than among O V languages if and only if X is a non-phrasal category and Y is a phrasal category.

Dryer points out himself, however, that the BDT fails to account for at least three correlation pairs: the order of verb and manner adverb (both of which are non-phrasal), the order of verb and subject (assuming that the subject NP (S) actually combines with the verb phrase (VP) we are dealing with two phrasal categories), and the order of affix and stem (which I will ignore here, but see Dryer 1992: 125–8 for discussion; see also Dryer 1992: 118 f. on ‘remaining complications’). Furthermore, since ‘the predictions of the BDT depend on one’s assumptions about a constituent structure’ (*ibid.* 109), Dryer subsequently reformulates his BDT twice to account for different subsets of all the correlation pairs. Thus, due to certain problems having to do with, for instance, the status of adjectives as a non-branching category and the constituent structure of NPs, Dryer proposed a (p.280) revised version and ultimately a more elegant ‘alternate version’ of the BDT.¹⁹ The main problem is, however, that all versions of the BDT (see note 19) are based on some form of a Chomskyan constituent tree, so that Dryer’s BDT only works for those who also accept certain features that are peculiar to the Chomskyan framework such as X-bar theory (also: X' or \overline{X} ; Jackendoff 1977), in which lexical and grammatical categories are treated alike (so that, for example, articles can also be regarded as ‘heads’ in the underlying structure; Abney 1987) and in which N-bar is a fully recursive phrasal category, whereas a noun (N) is a non-phrasal category. For reasons mentioned at the outset, the grammatical theory that I use for this book is Simon Dik’s *Functional Grammar* (FG). Below I will demonstrate how Dryer’s *Greenbergian word order correlations* can be accounted for in this theory.

9.3.3. Dryer’s correlation pairs and Functional Grammar

In Chapter 8 I wrote that in FG cross-linguistic ordering tendencies are described and explained in terms of a number of interacting and sometimes competing principles (Dik 1997: 395). In this section I will attempt to account for Dryer’s correlation pairs in terms of these general ordering principles.

9.3.3.1. The order of subject and object

Apart from the problem the pair verb-subject poses for Dryer’s BDT (see above), there is also the problem that, despite the correlation, SVO order is still the second most frequent order of subject, verb, and object among the world’s languages (Tomlin 1986). Furthermore Dryer (1992: 125) himself suggests that we may only be dealing with an ‘apparent correlation between the order of subject and verb and that of object and verb ... because of the rarity of OVS languages: it is because SVO is common while OVS is not

that subjects precede the verb more often in OV languages than they do in VO languages.' In my view the correlation in question is actually an instance of another, more general tendency, that is also apparent in Table 9.1, namely that the Subject precedes the Object. On the basis of a sample that contained over a thousand languages, Tomlin (1986) found that (p.281) languages were distributed over the six possible orderings of verb (V), subject (S), and object (O) in the following way:

(33)	SOV	SVO	VSO	VOS	OVS	OSV
	44.78%	41.79%	9.20%	2.99%	1.24%	0.00%

These data show that the object NP precedes the subject NP in only a very small minority of the world's languages. In FG (where Subject and Object function are interpreted in terms of a notion of 'perspective') this ordering tendency is regarded as a form of iconic linear ordering (Dik 1997: 406):²⁰

the Subject function marks that term which designates the entity from whose point of view the SoA [=State-of-Affairs, i.e. the kind of event defined by the clause, or rather: predication—JR] is presented; the Object function marks a secondary vantage point relative to the presentation of the SoA ... On that interpretation of Subject/Object, [the principle that says that Subject precedes Object] is certainly a 'natural' principle: it would be rather strange for a language to systematically prepose the Object to the Subject, where the Object marks the secondary, and the Subject the primary vantage point.

It seems therefore that there is only an apparent correlation between word order pairs verb-subject and verb-object, and that this correlation is essentially due to a strong universal tendency to refer to the subject entity before the object entity.

9.3.3.2. The position of relators

Several of Dryer's correlations have to do with what in FG are called 'relators' (coordinators, adpositions, case markers, subordinating elements):

1. the correlation between VO/OV and the order adposition-NP (*on + the table*) [V ... PREPOSITION NP ...] or [... NP POSTPOSITION ...V]
2. the correlation between VO/OV and the order complementizer—complement clause (*that + John is sick*): [V ... COMPLEMENTIZER complement] or [complement COMPLEMENTIZER ... V]
3. the correlation between VO/OV and the order adverbial subordinator—adverbial clause (*because + Bob has left*): [V ... SUBORDINATOR adv. clause] or [adv.clause SUBORDINATOR ... V]

The positional tendencies of these and other relators are captured by the *Relator Principle* (Dik 1997: 406), which says that relators have their preferred position (i) between their two relata, or (ii) at the periphery of the relatum with which they form one constituent (if they do so). As in the case of the principle concerning the relative order of Subject and Object, the *Relator Principle* can be seen as an instance of a more general iconic principle: since a relator marks some kind of (p.282) connection *between* elements, its most natural position is *between* its relata in the linear organization of the linguistic expression.

The correlation between VO/OV and the order of verb-PP (=verb-adpositional phrase (e.g. *slept + on the floor*) where the PP is regarded as an argument of the verb) is simply another case of V + argument NP. The only difference is that this argument requires the overt morphological expression of a relator (preposition or postposition). Although in the great majority of the languages adpositional arguments pattern just like object NPs, Dryer found that in a small number of OV-languages (apparently all non-rigid OV-languages from Africa and South America) PPs appear after the verb (OV and V-PP). He has only one instance of a language (from the SE Asia and Oceania area) that has the mirror pattern (VO and PP-V).

9.3.3.3. Dryer's correlation pairs and the Principle of Head Proximity

Seven correlation pairs display patterns that provide supporting evidence for the view that there is a cross-linguistic tendency to avoid having intervening material between heads of hierarchically ordered domains, i.e. for the *Principle of Head Proximity* (I have italicized the elements that apparently tend not to occur between heads of hierarchically ordered domains):

1. copular verbs tend to precede the sentence predicate (i.e. head of the clause) in VO-languages and vice versa in OV-languages.

(VO: *COPULAR VERB* PRED [NP] ... ('Mary's son is tall')

(*is COP* tall_{PRED} [son_N [[Mary]_G]_{NP}...])

OV:... [NP] PRED *COPULAR VERB* (... [[Mary]_G son_N]]_{NP} tall_{PRED} *is COP*)

2. the verb 'want' tends to precede the subordinate verb with which it is associated in VO and vice versa in OV-languages (i.e. most or all constituents in the subordinate clause do not occur in between the head of the matrix clause ('want') and the head of the complement clause (here: 'buy')):

('... wants to buy a car tomorrow')

VO: 'want' [V ...] ('want_v' [buy_v *car tomorrow*])

OV: [... V] 'want' ([*tomorrow car* buy_v] 'want_v')

3. auxiliary verbs (AUX) tend to precede the verb in VO-languages and vice versa in OV-languages:

('Mary's son has arrived ...')

VO: *AUX* V [NP]... (*has AUX* arrived_v... [son_N [[Mary]_G]_{NP}...])

OV:... [NP] V *AUX* (... [[Mary]_G son_N]]_{NP}... arrived_v *has AUX*)

(p.283)

4. negative auxiliary verbs tend to precede the verb in VO-languages and vice versa in OV-languages (see Dryer 1992: 101 for examples):

VO: *NEGATIVE AUXILIARY* V [NP]...

OV:... [NP] V *NEGATIVE AUXILIARY*

5. in languages in which question particles (typically used in yes/no questions) normally occur in a sentence-peripheral position, they tend to precede the verb in VO-languages and vice versa in OV-languages (see Dryer 1992: 102–3 for examples):

VO: *QUESTION PARTICLE* V [NP]...

OV:... [NP] V *QUESTION PARTICLE*

6. the possessor phrase (G) usually follows the noun in VO-languages and tends to precede the noun in OV-languages:

VO: V [N G]_{NP}...

OV:... [G N]_{NP} V

7. the relative clause commonly follows the noun in VO-languages and tends to precede the noun in OV-languages:

VO: V [N *REL*]_{NP}...

OV: ... [REL N]_{NP} V

Although the *Principle of Head Proximity* presented in the introduction to this chapter was initially formulated to explain word order tendencies inside the NP domain, we see here that its significance extends to ordering phenomena outside the domain of the NP as well (see also note 3). Nevertheless Dryer's data only partially confirm the second hypothesis in that they demonstrate that the *Principle of Head Proximity* only determines the position of embedded modifiers (G and Rel) relative to the head noun; I will return to this in section 9.4.4 below.

9.3.3.4. Remaining correlation pairs

Throughout this study I have distinguished between grammatical and lexical instances of noun modification, between embedded and non-embedded noun modifiers, and between elements that are an integral part of the NP and those that are not (such as appositions). The reason I take these distinctions into consideration is that they, in one way or another, bear on the syntactic properties of the element in question.²¹ To give an example, we saw that numerals are categorized as grammatical elements in some languages (e.g. English), but as lexical elements in other languages (e.g. Samoan, where numerals are regarded as verb; see Mosel and Hovdhaugen 1992: 318). If numerals are lexical elements, this not only (p.284) means that they are heads (i.e. sensitive to the *Principle of Head Proximity*), but also that they are part of an attributive NP or relative clause (i.e. an embedded modifier), whose syntax differs from that of grammatical expressions of cardinality (non-embedded modifiers). Syntactic differences between embedded and non-embedded modifiers are not only captured in the *Principle of Domain Integrity* (Chapter 8), but also in some of the principles formulated by Hawkins (1983). Consider, for example, his *Heaviness Serialization Principle*:

(34) *Heaviness Serialization Principle* (Hawkins 1983: 90–1) $\text{Rel} \geq_R \text{Gen} \geq_R \text{A}$
 $\geq_R \text{Dem}/\text{Num}$

'where " \geq_R " means "exhibits more or equal rightward positioning relative to the head noun across languages".' That is, heavier noun modifiers occur to the right [of the head noun] (Hawkins 1983: 90–1; see also Dik's LIPOC mentioned in Chapter 8).

Additionally Hawkins proposed the *Mobility Principle*, according to which demonstratives, numerals, and adjectives (i.e. non-embedded modifiers) can move around the head noun more easily than relative clauses and possessor NPs (Gen):

(35) *Mobility Principle* (Hawkins 1983: 93)

$$\left\{ \begin{array}{l} \text{Adj} \\ \text{Dem} \\ \text{Num} \end{array} \right\} \geq_M \left\{ \begin{array}{l} \text{Rel} \\ \text{Gen} \end{array} \right\}$$

‘where “ \geq_M ” means “exhibits greater or equal mobility from the adposition + NP serialization”.

Finally, we also saw that in quite a few languages that employ numeral classifiers, the numeral (plus classifier) is an appositive constituent, i.e. not an integral part of the noun phrase and therefore probably not sensitive to the principles that determine the position of modifiers that do form an integral part of the NP.

One of the problems of Dryer’s investigation (readily admitted by himself) is that in identifying certain categories he often only relied on semantic criteria. Consider, for example, what he writes about demonstratives, numerals, and adjectives, which, in combination with a noun, did not form a correlation pair according to his investigation (Dryer 1992: 96, 120, 122).

As discussed in Dryer 1988, there are many languages in which what I call adjectives are really verbs, and ‘adjectives’ modifying nouns are really just a kind of relative clause.

The standard practice in word order typology since Greenberg 1963 has been to identify different categories largely on the basis of semantic criteria. Underlying this practice is an implicit assumption that semantic categories like ‘number’ and ‘demonstrative’ correspond to universal grammatical categories or that any differences in the way in which these semantic categories are realized in particular grammars are somehow irrelevant to the correlations ... In at least some cases, then, failure to consider more specific grammatical properties in particular languages would obscure what is going on with the correlations.

(p.285) The situation with demonstratives is thus not unlike what we observed for numerals. While they may constitute a fairly well-defined category from a semantic point of view, their grammatical properties seem to vary from language to language.

Dryer’s remarks about the heterogeneous character of some of his categories (1992: 122) were made in the context of a discussion about certain non-correlation pairs (demonstrative-noun, numeral-noun, and adjective-noun), but similar things can be said about the four correlation pairs listed in Table 9.3 that have remained undiscussed so far: article + N’, plural word + N’, verb + manner adverb, and adjective + standard of comparison. Apart from the fact that two of these pairs involve N’ (a theoretical construct that does not exist outside Chomsky’s theory), the main problem with these correlation pairs is that they involve categories whose members are distributed over various word classes or construction types. Moreover, as was already mentioned earlier, apparently Dryer did not distinguish between embedded and non-embedded elements, or between integral and non-integral (e.g. apposed) noun modifiers. Below I will discuss each of these correlation pairs in some more detail.

9.3.3.4.1. Articles and plural words

In his study Dryer (1989a: 83) used two criteria to identify what he calls articles, but in my view the two groups of elements thus identified have very little in common. He counted as an article not only a word indicating (in)definiteness ‘or some related discourse notion’, but also a word serving as a noun phrase marker ‘in the sense that noun phrases in that language ... typically occur with one of the words in question’. This resulted in a group of elements that are formally identical or similar to a demonstrative

modifier or the numeral ‘one’ and another group whose members are quite distinct from a demonstrative modifier or the numeral ‘one’. The latter subgroup includes such varied items as noun (phrase) markers (as in the Fijian example below), noun classifiers (so that both *hune'* and *no'* count as articles in the Jacaltec example below), pronouns (such as *rá* in the example from Jicaltepec Mixtec), and items such as the ‘polyvalent lexeme’ *cái* in Vietnamese (Nguyễn Đình Hoà 1997: 180), which (it is true) some have called an article (or even ‘superarticle’; Bulteau 1953: 21) but whose exact meaning is far from clear. For example, it seems to serve as a classifier, as an emphasizing, individualizing, and particularizing element, as a modifier ‘to explain more clearly or to reinforce the meaning of the noun’ or ‘to attract the reader’s attention to the noun’, and it can be used to give the NP a pejorative connotation (Nguyễn Đình Hoà 1997: 179–80; cf. also Vũ Duy-Tù' 1983: 30, 32).

Boumaa Fijian (Dryer 1989a: 84—original example in Dixon 1988: 243)²²

(36)	sa	tau-ra	a	drano	o	Boumaa
	ASP	hold	ART	lake	ART	Boumaa
‘Boumaa held the lake’						

(p.286) Jacaltec (Dryer 1989a: 86—original example in Craig 1977: 137)

(37)	hune'	no'	txitam
	a	CLF	Pig
‘a pig’			

Jicaltepec Mixtec (Dryer 1989a: 93; original example in Bradley 1970: 79)

(38)	číká,	ča?	a	ňa	sa?ma	či?	rá	ahili
	thing-that	give	she	clothes	to	he	angel	
‘That’s why she gave the clothes to the angel’								

Vietnamese (Nguyễn Đình Hoà 1997: 180; original example in Lê Văn Lý 1960: 213)

(39)	Cái	con	ngựa	ấy	chạy	nhanh	thật
	CAI	CLF	horse	that	run	fast	real
‘That horse over there runs really fast’							

Similar things can be said about the plural words in that they, too, consist of elements belonging to various categories. This is also explicitly stated by Dryer (1989b: 870): ‘it is clear that the categorial status of plural words varies considerably in the languages of my sample. In general, plural words seem to fall into a variety of different categories: numerals, articles, a minor word class containing a few other words, or a one-word minor word class.’ Additionally the appendix that lists the forty-eight languages with

plural words give ‘miscellaneous’ and ‘unclear’ as categories. As in the case of Dryer’s category of articles, I doubt whether it is useful to make generalizing statements about such a mixed bag of elements which (in this case) only seem to have in common that they express some kind of plural notion. First of all, there appears to be an overlap between Dryer’s category of plural words and some of his other categories discussed above (such as articles). In other words, there is the possibility that the same element occurs in other correlation pairs as well.²³ Secondly, apart from the problem of heterogeneity in terms of word classes, it seems that at least some of Dryer’s plural words are elements that I regard as quantifiers (e.g. Vietnamese *nhຸ’ng*; see section 5.2.1.3 and Rijkhoff 2000) or collective markers (such as Hixkaryana *homo*; see section 4.2.1.2). Finally, since the category of ‘plural words’ also contains phrase-final third person plural pronouns (Dryer 1989b: 875–6) there is the distinct possibility that this category includes elements that are not an integral (p.287) part of the NP but phrasal modifiers (see section 2.2; cf. also Lehmann 1995 and Frajzyngier 1997: 199 f.).

Chamorro (Dryer 1989b: 876; original example in Topping 1973: 234)

(40)	a.	man-estidiante	siha
		PL-student	they
‘They are students’			
	b.	man-metgot i	estudiante siha
		PL-strong the	student PL
‘The students are strong’			

Next is an example from Mupun (which is not in Dryer’s sample), which illustrates my point even better; note that the ‘plural word’ *mo* is identical with the third person plural pronoun and must occur in NP-final position:

Mupun (West Chadic—Frajzyngier 1997: 200)

(41)	jirap	de	wuraj	mo
	girl	REL	tall	PL
‘tall girls’				

9.3.3.4.2. Manner adverbs

It is not entirely clear how Dryer identified members of his category of manner adverbs in the world’s languages, but considering his remarks about the way he identified members of other categories such as adjectives and numerals, it seems safe to assume that in the case of manner adverbs he also focused on semantic rather than formal properties. Recall, for example, that he stated that ‘there are many languages in which what I call adjectives are really verbs, and “adjectives” modifying nouns are really just a kind of relative clause’. Thus, since not every language has a separate class of manner adverbs either, one may expect his category of manner adverbs to contain various other kinds of manner expressions as well. At this point it is appropriate to briefly return to Hengeveld’s typology of parts-of-speech systems (see Chapter 1), the complete version of which also includes manner adverbs. Since manner adverbs are irrelevant in a study

of NPs referring to first order entities, I collapsed three types in Hengeveld's original typology (V-N-A/adv, V-N-A-adv, V-N-A) into one: V-N-A (type 3). Fig. 9.1 gives Hengeveld's complete classification, where 'adv' stands for 'manner adverb' (Hengeveld 1992b: 58). Recall that Hengeveld defined adjectives as predicates 'which, without further measures being taken, can be used as a modifier of a nominal head' and that he distinguished between

- languages without a class of adjectives (types 4–5)
- languages with a distinct class of adjectives (here: types 3b-3c)
- languages in which adjectives are not distinguished from members of one or more other word classes (here: types 1, 2, 3a).

(p.288)

	Type 1	V/N/A/adv		
Flexible	Type 2	V	N/A/adv	
	Type 3a	V	N	A/adv
	Type 3b	V	N	A
Rigid	Type 3c	V	N	A
	Type 4	V	N	—
	Type 5	V	—	—

Fig. 9.1. Parts-of-speech systems (based on Hengeveld 1992b)

Fig. 9.1 shows that we can make the same division regarding manner adverbs, which are defined as predicates 'which, without further measures being taken, can be used as a modifier of a non-nominal head':

- languages without a class of manner adverbs (types 3c, 4, 5)
- languages with a distinct class of manner adverbs (type 3b)
- languages in which manner adverbs are not distinguished from members of one or more other word classes (types 1, 2, 3a).

Here I will only give examples of the three subtypes of type 3, which were ignored in Chapter 1. Ngiti is a good example of a language of type 3a (Kutsch Lojenga 1994: 336):

There is no morphological nor a clear syntactic distinction between a class of adjectives and a class of adverbs in Ngiti. The functional term modifiers is therefore used ... to cover a fairly large grammatical class of words, containing about 150 items, which are neither nouns nor verbs and which all have a modifying function in relation to different constituents.

In the following examples, *isò* is first used adjectivally (to modify a noun) meaning ‘light (of weight)’, and then as a manner adverb meaning ‘easily, without effort’.

(p.289) Ngiti (Kutsch Lojenga 1992: 338)

(42) ngbángba ní-itdù	isò ànò
child	RSM-carry:PERF.PRES light load
‘the child carried a light load’	

(43) isò ngbángba ní-itdù	ànò
light child	RSM-carry:PERF.PRES load
‘the child carried a load easily’	

Wambon is a language that, apart from one or two exceptions, has no flexible or distinct class of adverbs (type 3c); instead Wambon employs medial verb constructions (de Vries 1989: 49):

The category of manner adverbs can be so marginal because Wambon prefers to use medial verbs as modifiers of other verbs in serial verb constructions in which the modifying verb immediately precedes the modified verb ... Very often the medial verbs specifying manner, are verbs which are derived from adjectives by - mo.

For example, in the next example the verb *matetmo* is derived from the adjective *matet* ‘good’

Wambon (de Vries 1989: 49)

(44) Jakhov-e matet-mo	ka-lembó?
they-CN good-SUPP.SS	go-3PL.PAST
‘Did they travel well?’	

Ngalakan, finally, belongs to type 3b, because it has a group of lexemes specifying ‘manner’ that can immediately be used to modify the verb, such as *yukalji* ² ‘thoroughly, forcefully, altogether, for good’, *yuča* ‘quickly’, *mapuy* ² ‘slowly’, *gamakun* ‘properly’ (Merlan 1983: 123).

9.3.3.4.3. Adjectives and comparative constructions

Apart from the fact that many languages simply lack a class of adjectives (section 4.3), there is also the problem that languages may employ rather different constructions to express comparisons of the kind ‘Mary is taller than Bill’. Stassen’s cross-linguistic investigation of the comparative construction (Stassen 1985, see also Stassen 1986) shows that there are at least six major construction types, two of which are exemplified below.

Duala uses two predicates: one expressing the adjectival notion, the other is the equivalent of the English verb ‘to exceed’ (hence Stassen calls it the ‘exceed construction’). This construction is the second largest type in Stassen’s 110-language sample and is used in e.g. Banda, Bari, Cambodian, Dagomba, Duala, Fulani, Gbeya, Hausa, Igbo, Jabem, Kirundi, Mandarin Chinese, Margi, Nguna, Swahili, Thai, Vietnamese, Wolof, Yagan, and Yoruba.

(p.290) Duala (Stassen 1986: 151)

(45) Nin ndabo e kolo buka nine

this house it big exceed that

‘This house is bigger than that’

Hixkaryana is one of the languages that employs what Stassen calls the ‘conjoined comparative’, which typically involves ‘the adversative coordination of two clauses, which exhibit a structural parallelism’ (Stassen 1986: 157). Other languages in Stassen’s sample employing this construction are Abipon, Cayapo, Dakota, Ekagi, Gumbainggar, Kobon, Mangarayi, Maori, Menomini, Miskito, Mixtec, Motu, Monumbo, Classical Nahuatl, Pala, Samoan, Shipibo, Sika, and Yavapai.

Hixkaryana (Stassen 1986: 154)

(46) Kaw-ohra naha Waraka, kaw naha Kaywerye

tall-Not he:is Waraka tall he:is Kaywerye

‘Kaywerye is taller than Waraka’

Also, quite a few languages have more than one way to form a comparative construction and some use a ‘mixed comparative’, that is, ‘a comparative construction in which the fundamental features of two typological variants are combined’ (Stassen 1986: 147).

9.3.3.5. Conclusion

According to the second hypothesis that was formulated on the basis of the strong version of the *Principle of Head Proximity* all noun modifiers (but especially embedded noun modifiers) tend to avoid appearing in between the head of the noun phrase N and the head of the clause V. To the extent that Dryer’s investigation gives us a reliable picture of correlations between the order of various noun-modifier combinations and word order patterns at the sentence level, we may conclude that this hypotheses is confirmed with respect to embedded noun modifiers (possessor NPs and relative clauses), but that the *Principle of Head Proximity* does not seem to affect the position of demonstratives, numerals, and adjectives relative to the head noun. Presumably, in

the NP ‘head proximity’ is only seriously obstructed by relatively long and complex modifiers, such as possessor NPs and relative clauses, which furthermore have the potential to contain other such embedded modifiers.

- (47) Preferred position of the possessor NP (G) and the relative clause (Rel) relative to the head noun:

V [N[G/Rel]]NP...	(VO/V-initial)
... [[G/Rel]N]NP V	(OV/V-final)

- (48) Non-preferred:

V[[G/Rel]N]NP...	(VO/V-initial)
... [N[G/Rel]]NP V	(OV/V-final)

(p.291) Another possible reason why only embedded modifiers confirm the second hypothesis is that they are also referring expressions, which, in terms of processing ease, may present more of an obstacle in the NP than non-referring material. In the case of a possessor NP the referent is a first order entity, but in the case of a relative clause the referent is an event (action, process, etc.) involving one or more participating parties. Should an embedded domain occur on the wrong side of the noun in terms of the *Principle of Head Proximity* (as in (48)), this would not only increase the distance between N and V, but it would also force the processor to construe or identify one or more referents that only have an indirect relationship with the main verb.²⁴ For example, in ‘(I) bought Mary’s book’, the human processor first has to deal with Mary, an entity that has no immediate connection with the buying event as such, and only then with the book, the entity that actually has a direct relationship with the verb (I bought the book, not Mary; she is only mentioned to ‘ground’ the book; cf. note 24). Conversely, in a V-final language the language user would first encounter the referent of the matrix NP (‘book’), but before the human parser reaches the main predicate it has to deal with a subordinate domain, referring to the possessor of the book (‘Mary’), and only then does it encounter the head of the superordinate domain (‘buy’).

Before drawing any conclusions, however, we must remember that Dryer did not distinguish between lexical or grammatical instances of modification, nor did he take into account the syntactic relation between the noun and the (real or apparent) modifier (embedded vs. non-embedded; integrated vs. not integrated). This leaves open the possibility (to use Dryer’s words; see above) that ‘the failure to consider more specific grammatical properties in particular languages’ has indeed obscured what is going on with the correlations. So before we accept that across languages the position of the demonstrative, the numeral, and the adjective is not influenced by the *Principle of Head Proximity*, we will consider the data for the languages in the fifty-two-language sample, bearing in mind of course that this sample is a *variety sample*, which is not designed to yield statistically significant correlations or distribution patterns (Rijkhoff and Bakker 1998). First, however, we will discuss another major proposal to explain Greenbergian word order correlations (among others): Hawkins’s principle of *Early Immediate Constituents* (EIC).

9.3.4. Hawkins's principle of Early Immediate Constituents (EIC)

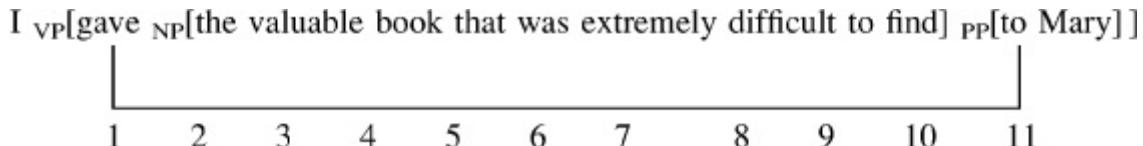
In 1990 Hawkins proposed 'a theory of syntactic processing' called EIC (*Early Immediate Constituents*) which says that 'words and constituents occur in the orders they do so that syntactic groupings and their immediate constituents (ICs) (p.292) can be recognized (and produced) as rapidly and efficiently as possible in language performance' (Hawkins 1994: 57, 106):²⁵

Early Immediate Constituents (Hawkins 1994: 77)

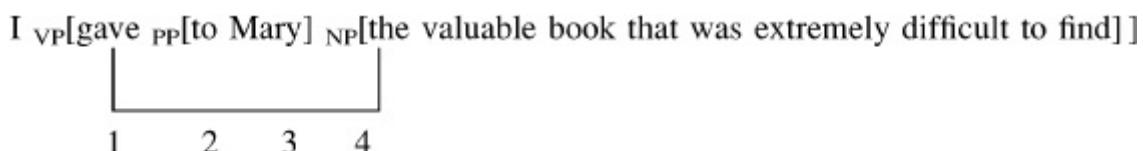
'The human parser prefers linear orders that maximize the IC-to-non-IC ratios of constituent recognition domains.'

As an illustration of the EIC principle, Hawkins provides the following example of 'Heavy NP Shift' from English (my numbering—JR):

(49)



(50)



Example (50) provides a more rapid presentation of the three ICs of the VP (V, NP, and PP) than (49). The verb *gave* is the first IC of the VP in both examples and signals to the parser that a VP should be constructed. The PP is a two-word IC here. Its positioning to the left of the lengthy NP in (50) makes it possible for all three daughter ICs to be recognized within a short viewing window, since the NP can be recognized on the basis of the determiner *the*, occurring in leftmost position within this NP. In (49), on the other hand, the viewing window extends all the way from *gave* to the preposition *to*, which constructs the PP, and the heaviness of the intervening NP delays access to this third IC. Of the 12 total words dominated by this VP, therefore, 11 need to be examined for IC recognition in (49), whereas just four suffice in (50). (Hawkins 1994: 57)

A little further he adds:

I believe that this preference for (50) is not something that is unique to Heavy NP Shift, or even to other rightward-moving rearrangements in English and other languages. Rather, what we are seeing here is just one manifestation of an extremely general principle of ordering that holds, I shall argue, not only for rearrangement phenomena but for basic word orders as well, not only in English but in all languages, and not only throughout the grammar, but throughout language performance as well. (Hawkins 1994: 58)

Hawkins claims that the EIC principle explains many ordering conventions of various grammatical theories, including Dik's Functional Grammar (Hawkins 1994: 58, 120; see also Hawkins 1997: 750–1). As a matter of fact, he states that (p.293) the *Principle of*

Head Proximity, as well as the *Principle of Domain Integrity*, among others, ‘follow from EIC (*ibid.*)’.²⁶

In this section I will compare EIC with the *Principle of Head Proximity* and show that predictions based on EIC largely overlap with those based on the *Principle of Head Proximity*, which was originally proposed in my MA thesis and published in a revised version two years later (Rijkhoff 1986). First, however, it is important to make sure that we are making a fair comparison. Although Hawkins (1994: 72) writes that it is his strategy to be conservative in his theoretical assumptions and to maximize the common ground, it is clear that a Chomskyan-style generative framework serves as a general theoretical background (Hawkins 1994: 411). For example, apart from the fact that underlying structures are consistently presented in the form of Chomskyan constituent trees, he assumes that a sentence contains a Verb Phrase {V, NP, PP} and that a preposition is a constructing category (1994: 120, 328). These are theory specific assumptions that are not necessarily shared by other general theories of grammar. Obviously Dik’s theory of Functional Grammar differs fundamentally from any version of Chomsky’s Generative Grammar and some of the differences between the two approaches have to be made explicit so as to be able to properly assess Hawkins’s statement that a number of FG ordering principles are reducible to his EIC.

9.3.4.1. EIC and ordering principles in Functional Grammar

Hawkins mentions four FG ordering principles that ‘appear to be reducible to one and the same principle’ (Hawkins 1994: 120). Since I believe that EIC largely amounts to the same as the *Principle of Head Proximity*, I will first argue that the three other principles he mentions, the *Principle of Domain Integrity*, *LIPOC* (*Language Preferred Order of Constituents*), and the *Relator Principle* (which were all introduced in Chapter 8), cannot be reduced to EIC.²⁷

9.3.4.1.1. The Principle of Domain Integrity and EIC

I have already mentioned in Chapter 8 that the *Principle of Domain Integrity* accounts for the fact that discontinuity is a marked phenomenon and can be regarded as the FG equivalent of ‘constituency’ in other theories:

Constituents prefer to remain within the boundaries of their proper domain; constituents of matrix domains prefer not to be interrupted by embedded domains.

Constituency is basically concerned with the observation that linguistic elements that (semantically) form part of a larger whole normally occur together (p.294) (syntactically) in the actual verbal expression. The reason why this phenomenon can be taken for granted in Generative Grammar (but not in Functional Grammar) is that constituency is simply ‘generated’ by structure building rules in the grammar. Therefore, it would be too simple to say that constituency is ‘reducible to’ or ‘subsumed by’ EIC (cf. Hawkins 1994: 118–20). What EIC does help to explain is that discontinuity is sometimes preferred, as when it improves efficient language processing. The idea that discontinuity may serve as a strategy to facilitate language processing for the human parser was also discussed in the context of head proximity (Rijkhoff 1986: 113).

9.3.4.1.2. LIPOC and EIC

According to Hawkins (1994: 120) ‘both LIPOC and Head Proximity follow from EIC (recall that Head Proximity is momentarily ignored). LIPOC (*Language Preferred Order of Constituents*) has been discussed in Chapter 8 and reads as follows (Dik 1997: 411):

(51) Other things being equal, constituents prefer to be placed in an order of increasing complexity, where complexity of constituents is defined as follows:

- (i) clitic < pronoun < NP < adpositional phrase < subordinate clause;
- (ii) for any category X: X < X coordinator X;
- (iii) for any categories X and Y: X < X subordinator Y.

Dik divides constituent ordering principles into three levels of abstraction. At the most abstract level we find general principles such as the *Principle of Head Proximity* and the *Principle of Domain Integrity*. At the intermediate level we find more specific reflections of these abstract principles (such as LIPOC), and the most concrete manifestations of constituent ordering principles are the language specific ordering rules. As Dik (1977: 399) put it: ‘The idea is that the general principles are reflected in the specific principles, and the specific principles in the actual constituent ordering patterns of individual languages.’

LIPOC is a specific instance of the more general *Principle of Increasing Complexity*, which says that there is ‘a preference for ordering constituents in an order of increasing complexity’ (Dik 1997: 404). The formulation of this principle was inspired by Behaghel’s *Gesetz der wachsenden Glieder* (Behaghel 1932: 6), which postulates a general preference for shorter elements to precede longer elements. Hawkins claims that the short-before-long ordering preference (as captured in Behaghel’s law), as well as the mirror pattern (long-before-short, in OV-languages), follow from EIC, but when he discusses some typological asymmetries (Hawkins 1994: 325–6) it becomes clear that EIC does not explain the rightward skewing for sentential direct objects (they typically occur after the verb in VO-languages, but are attested on either side of the verb in OV-languages; cf. Hawkins 1994: 264, 325). According to Hawkins this asymmetry is not due to the fact that sentences containing long and complex constituents are on the whole easier to process when these complex constituents appear after all (p.295) other (shorter, simpler) constituents of the sentence, but rather occurs because in an OV-language the human parser has to delay the decision about the main-or-subordinate status of the clause until the appearance of some disambiguating sign, such as a (clause-final) subordinating conjunction (note that EIC prefers clause-final complementizer in an OV-language; Hawkins 1994: 324–5). For that reason Hawkins introduces a principle that interacts with EIC and that states that there is a universal parsing preference for ‘immediate matrix disambiguation’ (IMD; Hawkins 1994: 324, also Hawkins 1997: 759), i.e. the presence of some element early in the structure that unambiguously indicates whether that structure is a main or a subordinate clause. The *Principle of Matrix Disambiguation* ‘predicts, in conjunction with EIC, a number of left-right assymmetries involving the positioning of sentential nodes across languages’ (Hawkins 1997: 759). The same principle also accounts for the rightward skewing for relative clauses, which almost exclusively occur after the noun in VO-languages, but are found on either side of the noun in OV-languages (Hawkins 1994: 265, 327).

The stipulation of the principle of *Immediate Matrix Disambiguation* as one of the principles that interact with EIC means that Hawkins, too, needs a separate principle to account for typological left-right asymmetries. In sum, the rightward skewing for sentential direct objects and relative clauses captured by LIPOC is not reducible to EIC; they are simply captured by another principle that Hawkins proposes: the principle of *Immediate Matrix Disambiguation*.²⁸

9.3.4.1.3. The Relator Principle and EIC

The *Relator Principle* (repeated below for convenience) is one of the *specific* FG ordering principles, i.e. a reflection of a more general (iconic) ordering principle, and refers to the position of adpositions and case markers, as well as coordinating and subordinating conjunctions (Dik 1997: 407):

(52) Relators have their preferred position

- (i) in between their two relata;
- (ii) at the periphery of the relatum with which they form one constituent (if they do so).

The elements that together make up the relator category in Functional Grammar are not treated in a unified fashion in Generative Grammar and this is also the approach adopted by Hawkins, who regards the adposition as a constructing category which forms an integral part of the prepositional or postpositional phrase. In Functional Grammar relators are not regarded as ‘constructing categories’, (p.296) which means among other things that the syntactic properties of adpositions and other elements with a similar (relator) function, such as subordinating and coordinating conjunctions, must be accounted for by a separate principle: the *Relator Principle* (note that Generative Grammar has to give a separate account for relators that are not ‘constructing categories’). In short, the *Relator Principle* does not ‘follow from EIC (and neither do the *Principle of Domain Integrity* and *LIPOC*). It is the particular set of background assumptions adopted by Hawkins rather than the principle of *EIC* that permits him to include the position of (certain) relators in EIC predictions.

9.3.4.2. The Principle of Head Proximity and EIC

To the extent that we can ignore differences due to theoretical background assumptions, predictions based on EIC essentially boil down to the same general ordering preferences that can be predicted on the basis of the *Principle of Head Proximity*, both for the sentence and for the NP. Furthermore, both EIC (Hawkins 1994) and the *Principle of Head Proximity* (Rijkhoff 1986) mention discontinuity and extraposition as strategies to increase EIC and Head Proximity ratings, respectively.

This does not imply, however, that EIC and the *Principle of Head Proximity* are variations of a single, more abstract ordering principle. Although in either case the notion of *distance* between certain elements in the clause plays a central role, EIC emphasizes the role of constituent recognition (e.g. VP, NP, PP), whereas the *Principle of Head Proximity* focuses on the importance of establishing relations between heads of hierarchically ordered domains (e.g. clause, NP). In this respect EIC and the *Principle of Head Proximity* ultimately go back to two different parsing (construction) strategies for English that were already discussed in Clark and Clark (1977: 59–62; note that Strategy 1 is originally due to Kimball 1973).

(53) Strategy 1: Whenever you find a function word, begin a new constituent larger than that word.

Some variations of Strategy 1 for English are (Clark and Clark 1977: 59):

Strategy 1a. Determiners and quantifiers

Whenever you find a determiner (*a, an, the*) or quantifier (*some, all, many, two, six, etc.*), begin a new noun phrase (NP).

Strategy 1b. Prepositions

Whenever you find a preposition (*to, at, in, etc.*), begin a new prepositional phrase (PP).

Strategy 2 has already been mentioned in section 9.1 and reads:

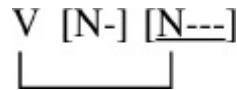
(54) Strategy 2: After identifying the beginning of a constituent, look for content words appropriate to that type of constituent.

The content word that is most appropriate for the noun phrase is of course the (head) noun.

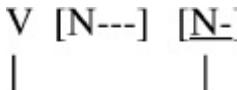
(p.297) The major prediction of EIC is that in an OV-language there is a preference for long-before-short, whereas the order short-before-long is preferred in a VO-language (Hawkins 1994: 118), as illustrated above in the introduction to section 9.3.4. The same ordering pattern is preferred by the *Principle of Head Proximity*, since this gives a better overall rating in terms of head proximity (a hyphen ‘-’ symbolizes a modifier category in the noun phrase):

(55)

a. VO (preferred)



b. VO (*not preferred*)



(56)

a. VO (preferred) [---N] [-N] V

b. VO (*not preferred*) [-N] [--N] V

In the non-preferred patterns the distance between the verb and the head of the underscored NP is longer (by two modifier categories, e.g. dem and A) than in the preferred pattern, i.e. with the order of the NPs reversed. Hawkins's data confirm that this is the preferred order of subordinate domains in VO- and OV-languages, even when we take into account the fact that his domains ('constituents') may include an adposition; consider, for instance, his remarks on the order of object NP and adpositional NP in English (Hawkins 1994: 89; see also p. 99):

[For example], the order [NP PP] is significantly more frequent than the reverse in the English VP ... direct object NPs in English are significantly shorter than

PPs: the average size of this NP in my text counts is 2.5 words, compared with 4.6 words for PP.

In other words, even without the preposition the average English adpositional NP in [V NP PP] is longer ($4.6 - 1 = 3.6$ words) than the object NP (2.5 words), yielding an ordering pattern that is most preferred in terms of the *Principle of Head Proximity*.

9.3.4.2.1. Head Proximity, EIC, and Dryer's correlation pairs

Both Hawkins and I have used the results from Dryer's (1992) large-scale cross-linguistic study on correlation pairs to investigate to what extent Dryer's findings coincide with preferred ordering patterns in terms of EIC and Head Proximity, respectively. In Table 9.4, I have listed Dryer's correlation pairs once again (see also section 9.3.2) and indicated which pairs are in accordance with the *Principle of Head Proximity* (they are underscored) and which pairs are mentioned in connection with the principle of EIC (they are in italic print; Hawkins 1994: 255–65); for reasons discussed above I will ignore the pairs 'adposition-NP' and 'verb-PP'.²⁹

(p.298)

Table 9.4. Dryer's correlation pairs

Verb patterners	Object patterners	Example
verb	subject	(there) entered + a tall man
adposition	NP	on + the table
<u>copula verb</u>	<u>predicate</u>	<u>is+a teacher</u>
<u>'want'</u>	<u>VP</u>	<u>wants+to see Mary</u>
<u>tense/aspect aux. verb</u>	<u>VP</u>	<u>has+eaten dinner</u>
<u>negative auxiliary</u>	<u>VP</u>	(see Dryer 1992: 101)
complementizer	S	that + John is sick
<u>question particle</u>	<u>S</u>	(see Dryer 1992: 102–3)
adverbial subordinator	<u>S</u>	because + Bob has left
article	N'	the + tall man
plural word	N'	(see Dryer 1992: 104–5)
<u>noun</u>	<u>genitive</u>	<u>father+of John</u>
<u>noun</u>	<u>relative clause</u>	<u>movies+that we saw</u>
adjective	standard of comparison	taller + than Bob
verb	PP	slept + on the floor
verb	manner adverb	ran + slowly

Source: Dryer (1992: 108).

Apart from the correlation pair 'question_particle-S', which Hawkins does not mention, it turns out that EIC and the *Principle of Head Proximity* account for the same word

order data.

9.3.4.2.2. The order of adjective and relative clause

Hawkins (1994: 269–74) also discusses the relative order of adjective and relative clause in the noun phrase and here, too, the basic predictions overlap in that both EIC and the *Principle of Head Proximity* predict that the relative clause (a longer domain) does not occur between the head noun and the adjective (section 9.2).³⁰ Of course, this is also the preferred order according to the *Principle of Domain Integrity* (Chapter 8). Thus, abstracting away from the OV/VO distinction and the tendency for relative clauses and other complex structures to appear later in the clause (section 9.3.4.1.2 and Rijkhoff 1997), the adjective and the head noun (p.299) are always adjacent in the preferred pattern:³¹

- (57) preferred: a. Rel A N b. A N Rel
c. Rel N A d. N A Rel

By contrast, orderings in which the relative clause appears between the adjective and the head noun are not preferred:

- (58) not preferred: a. A Rel N b. N Rel A

To my knowledge the only counter-example is Nenets (Uralic-Yukaghir). Assuming that this language actually has a true class of adjectives, the relative clause is said to appear between the adjective and the noun in the ‘normal’ order [dem num A Rel N]. There is, however, considerable variation in the relative order of noun modifiers (Tapani Salminen, personal communication); in fact only the NP-initial position of the demonstrative is completely fixed (cf. Rijkhoff 1997: 368). Hawkins (1994: 270, 388) mentions two other languages that are deemed to have the relative clause between the adjective in a non-basic pattern: Lahu and Mandarin Chinese. Recall, however, that I have argued in section 4.3.3 that Mandarin Chinese does not have a distinct class of adjectives; instead adjectival notions are expressed by verbs.

9.3.4.3. Conclusion

In this section I have argued that, to the extent that we can ignore differences that can be attributed to distinct theoretical background assumptions, EIC and the *Principle of Head Proximity* appear to cover more or less the same ground. That is to say, in the most important areas Hawkins’s parsing theory of word order regularities makes essentially the same predictions as the *Principle of Head Proximity*: (i) a short-before-long ordering preference in VO-languages and a long-before-short preference in OV-languages (Hawkins 1994: 118), and (ii) a strong tendency to avoid the occurrence of adnominal modifiers (notably the possessor NP and the relative clause) between the head of the clause and the head of the NP.

9.4. Basic word order and the position of noun modifiers in the fifty-two-language sample

For ease of comparison with Dryer's findings I have reclassified the languages in the fifty-two-language sample (see Table 8.1) in terms of the two-way VO/OV distinction, but also added as a separate type a group in which the VO pattern is (p.300)

Table 9.5. Word order data of sample languages (not included: Cayuga, Etruscan, Meroitic)

Language	dem	num	A	g	G	Rel	wo	adp/case
Babungo	+	[+]	{+}	+	+	+	VO	pr
Berbice Dutch Creole	+	-	-	±	±	+	VO	pr/po
Guaraní	-	-	+	px	-	+	VO	po
Gude	+b	+	{-}	sx	+	+	VO	pr
Hixkaryana	-			px	-		VO	po
Hmong Njua	+	-	+	-	-	+	VO	pr
Krongo	+	[+]	[+]	+	+	+	VO	pr/po
Lango	[+]b	[+]	[+]	sx	+	+	VO	pr
Ngiti	-	-	-	-	-	-	VO	po
Nung	+	-	[+]	+	+	+	VO	pr
Pipil	-	[-]?	{-}?	ax?	+	+	VO	pr/po
Samoan	-	[+]	+	sx?	+	+	VO	pr
Tsou	+	-	[-]	sx	+	-	VO	pr
Vietnamese	+	-	[+]	+	+	+	VO	pr
Chukchi	-b	[-]b	{±}b	px	-	?	VO/OV	po
Dutch	-	-	-	±	±	+	VO/OV	pr/po
Ket	-	-	-	px	-	+	VO/OV	po
Kisi	+	+	{+}	+	+	+	VO/OV	pr/po
Nunggubuyu	[±]?	[±]?	[±]	±?	±?	[adj]	VO/OV	po
Abkhaz	-	±b	+b	px	-	-	OV	po
Alamblak	-	(-)	-	-	-	±	OV	po
Basque	+	-b	+	-	-	-	OV	po
Bambara	-	[+]	{+}	-	-	[correl]	OV	po
Bukiyip	-	-	-	-	-	-	OV	po
Burmese	-	[+]	[+]	-	-	-	OV	po
Burushaski	-	-	-	-	-	-	OV	po
Galela	+	+	[+]	-	-	+	OV	pr/po
Georgian	-	-	-	-	-	+	OV	po
Greenlandic, West	[±]?	[+]	[+]	sx	-	+	OV	po
Hittite	-	-	-	-	-	[correl]	OV	po
Hungarian	-	-	-	sx	-	+	OV	po
Hurrian	+		-	sx	-	?	OV	po

	-	±	+	±	±	+	OV	po
Ika	-	±	+	±	±	+		
Kayardild	-	[−]	-	-	-	-	OV	po
Koasati	-		[+]	px	-	-	OV	po
Korean	-	[+]	[−]	-	-	-	OV	po
Mandarin Chinese	-	-	[−]	-	-	-	OV	pr/po
Nahali	-	-	-	-	-	-	OV	po
Nama Hottentot	-	-	-	-	-	-	OV	po
Nasioi	±	+	+	-	-	-	OV	po
Ngalakan	[−]?	-	+	sx	±?	[adj]	OV	po
Nivkh	-b	±b	[−]b	px	-	-	OV	po
Oromo	+	+	{+}	+	+	+	OV	po
Quechua, Imbabura	-	-	-	-	-	-	OV	po
Sarcee	-	[−]	{+}b	px	-	+	OV	po
Sumerian	+b	+	+	sx	+	+	OV	po
Tamil	-	-	{−}	-	-	-	OV	po
Turkish	-	-	-	sx	-	-	OV	po
Wambon	-	±	+	px	-	-	OV	po

(p.301) just as common as the OV pattern (Chukchi, Dutch, Ket, Kisi, Nunggubuyu) (see Table 9.5). The group of VO-languages are all those languages that were classified as V-initial or V-second (V--- or -V--); V-final languages, be they rigid or non-rigid V-final (---V or --V-), are now classified as OV-languages.

The minus symbol, i.e. ‘−’, indicates that the basic order is modifier-noun, a ‘+’ that the basic order is noun-modifier; a ‘=b’ means that both orders are more or less equally basic. Since I have ignored marginal orderings or patterns due to emphasis, there are some differences from the data in tables presented in Chapters 4–6. Note furthermore that bound modifiers are not taken into account, so that bare +, −, and ± only refer to the position(s) of a free modifier. In the case of possessor pronouns, however, I have specified whether they are expressed as prefixes (px), suffixes (sx), or both (ax), since this is relevant for certain universals that will be discussed in section 9.4.2.2. A question mark indicates uncertainty or lack of information.

Since this study is mainly concerned with simple, integral NPs, I have used square brackets to indicate that the demonstrative, numeral, or adjective is regarded as a non-integrated modifier or an embedded domain (thus creating a complex NP); a ‘b’ indicates that some or all members of that particular modifier category may or must be expressed as a bound form (see Chapters 4, 5, and 6 for details).³² Additionally I have used round and curly brackets. Round brackets are used in the numeral column (num) and signal that only some (rather than all) numerals are nouns or verbs (section 5.3). The curly brackets, which are only found in the adjective column (A), indicate that we are at best dealing with a minor class of true adjectives (section 4.3.2).

As to possessor NPs, so as not to make matters unnecessarily complicated I have not taken into account that in quite a few languages a possessor NP (G) requires a co-

referential pronominal (marked as xr or XR in Table 6.3), although (p.302) one could argue that such cross-referenced possessor NPs are only indirectly modifying the head noun, i.e. through the possessor pronoun. Recall, finally, that I only refer to pre- and postnominal relative clauses inside the NP, thus ignoring the head-internal relative construction that is employed in Ika, Imbabura Que-chua, and Wambon, the correlative construction in Bambara and Hittite, and the adjoined relative construction found in Ngalakan and Nunggubuyu. In all cases I refer to Chapters 4, 5, and 6 for the relevant details.

Since I have used a variety sample rather than a probability sample, I will refrain from making any statistical claims in my discussion of syntactic patterns in the sections below. Instead a preferred pattern (in terms of the *Principle of Head Proximity*) will only count as supporting evidence for the *Principle of Head Proximity* if it occurs at least twice as often as the non-preferred counterpart in Fig. 9.2 (data supporting the *Principle of Head Proximity* (=PHP) are indicated in bold print).

9.4.1. Demonstratives, numerals, and adjectives

Fig. 9.2 shows that the sample contains no VO-languages with doubling of the demonstrative, numeral, or adjective in the basic word order.³³ In fact the only case of doubling in the VO-languages concerns Berbice Dutch Creole, an SVO language which has doubling of the possessor NP (GN and NG, as in Dutch). All other cases of doubling are confined to languages with OV and VO/OV order.

The position of the demonstrative relative to the noun did not correlate with the position of object relative to the verb in Dryer's sample, which contained over 600 languages (Dryer 1991: 462). Languages in my much smaller sample, however, do show a tendency to adhere to a consistent head-initial (VO and Ndem) or head-final pattern (OV and demN) here. This holds especially for the OV-languages: for each OV-language with the order [N dem], there are four with the order [dem N]. Dryer (1992: 96) has pointed out that relatively many cases of languages with VO and Ndem in his sample are from the African continent and the same holds true in my sample: of the eight VO-languages with the order [N dem], five are spoken in Africa: Babungo, Gude, Krongo, Lango, and Ngiti.

Greenberg (1975: 29) has already noted that across languages numerals seem to prefer prenominal placement and this global tendency also shows up in my sample. According to Dryer (1992: 118), however, 'the two orders of numeral and noun are equally common among OV languages', whereas 'VO languages exhibit a strong tendency to be NumN'. Only the VO-languages that are spoken in Africa have a strong tendency for the numeral to follow the noun in VO-languages in Dryer's sample (recall that numerals are discussed under the heading of 'remaining complications' in Dryer 1992).

(p.303)

BWO	49	Rel	G	A	num	dem
PHP :		11 : NRel	9 : NG	4 : NA →1 : N{A}	1 : Nnum	8 : Ndem
VO	14	2 : RelN	4 : GN 1 : GNG	4 : AN →2 : [AJ]N	8 : numN	4 : demN
				4 : N[A]	4 : N[num]	1 : N[dem]
				1 : [AJ]N	1 : [num]N	
		1 : 0		1 : 0		1 : 0
VO/OV	5	3 : NRel	1 : NG 2 : GN 2 : GNG	2 : AN 1 : N{A} 1 : {A}N[A]	1 : Nnum 2 : numN	1 : Ndem 3 : demN
		1 : adjoined 1 : ?		1 : [AJ]N[A]	1 : [num]N 1 : [num]N[num]	1 : [dem]N[dem]
PHP:		8 : NRel 13 : RelN 1 : RelNRel	2 : NG 26 : GN 2 : GNG	11 : NA →4 : N{A} 13 : AN →1 : {A}N	5 : Nnum →1 : N(num) 14 : numN →1 : (num)N 4 : numNnum	5 : Ndem 22 : demN 1 : demNdem
OV	30	2 : correl 1 : adjoined		3 : N[A] 3 : [AJ]N	3 : N[num] 2 : [num]N	1 : [dem]N 1 : [dem]N[dem]

Fig. 9.2. Noun modifiers in the fifty-two-language sample and the VO/OV distinction (Cayuga, Etruscan and Meroitic are not included)

(p.304) Although I have only considered languages with a distinct class of adjectives, the result is the same as in Dryer's investigation: in either sample there is no correlation between the order of adjective-noun and the order of verb and object NP.

9.4.1.1. Greenberg's Universal 18

Greenberg's Universal 18 deals with the position of demonstrative, numeral, and adjective relative to the noun. It reads as follows (Greenberg 1966a: 86):

Universal 18. When the descriptive adjective precedes the noun, the demonstrative and the numeral, with overwhelmingly more than chance frequency, do likewise.

The reason this universal was formulated as a statistical implication is that Greenberg was aware of a very small number of exceptional languages (such as Efik and Zande—both Niger-Congo; Greenberg 1989: 113; see also Dryer 1988: 208). Table 9.6 shows that this universal is confirmed by data from my sample. Berbice Dutch Creole and Hurrian have prenominal adjectives but demonstratives follow the noun. In Gude both numerals and demonstratives appear after the noun, but this language only has a minor class of true adjectives, which normally precede the noun (when they follow they occur with a suffix that strongly

Table 9.6. The relation between the position of adjective, demonstrative, and numeral

49	demN and numN	Ndem and/or Nnum
Igs.		
AN	14 lgs.: Bukiyp, Burushaski, Dutch, Georgian, Hittite, Hungarian, Kayardild ([num]N), Ket, Nahali, Nama Hottentot, Ngiti, Imbabura Quechua, Turkish	2 lgs.: Berbice Dutch Creole, Hurrian
NA		10 lgs.: Abkhaz, Ika, Basque, Guaraní, Hmong Njua, Nasioi, Ngalakan, Samoan, Sumerian, Wambon
ANA		1 lge.: Alamblak
{A}N	2 lgs.: Pipil?, Tamil	1 lge.: Gude
N{A}	1 lge.: Sarcee?	4 lgs.: Babungo, Bambara, Kisi, Oromo
{A}N{A}	1 lge.: Chukchi	
[A]N	1 lge.: Mandarin Chinese	3 lgs.: Korean, Nivkh, Tsou
N[A]		7 lgs.: Burmese, Galela, Krongo, Lango, Nung, Vietnamese, West Greenlandic
[A]N[A]		1 lge.: Nunggubuyu?
?	1 lge.: Hixkaryana	

(p.305) resembles the definiteness marker; see section 4.3.2). Since in Nivkh, Korean, and Tsou (section 4.3.3) adjectival notions are expressed by verbs, they do not go against the general tendency. I have no languages with NA and demN and/or numN in my sample, but there are numerous examples of such languages outside this sample, such as the Romance languages (cf. Hawkins 1983: 71 f.).

9.4.2. Possessive modifiers

As in Dryer's sample, my data also indicate that there is a correlation between VO/OV order and the position of the possessor NP relative to the noun. In the VO-languages

there are about twice as many languages with NG order as with the order GN. In the OV-languages the figures are even more convincing. My sample has twenty-six OV-languages with GN order and only two with the order NG, namely Oromo and Sumerian (cf. also Hawkins 1983: 285), but note that in both languages the possessor NP can also precede the noun (see section 6.3.2.2 for details):

Oromo (Stroomer 1987: 178, 180)

(59) gogaa k'urfummii guddiyoo aa	
skin	fish
big LIN = GEN	
'the skin of a big fish'	

(60) sareeni tana k'eesa išii aa c'aafuu	
dog	this paw her LIN = GEN dirt(y)
'the paw of this dog was dirty'	

Sumerian (Thomsen 1984: 91)

(61) é lugal-ak	
house	king-GEN
'the house of the king'	

(62) lugal-ak é-ani	
king-GEN	house-his
'the house of the king'	

9.4.2.1. Possessor pronouns

Dryer (1992) does not mention the position of free possessor pronouns, but my data suggest that they too tend to appear in the preferred position in terms of the *Principle of Head Proximity*, particularly in OV-languages (sx = suffix; px= prefix). The only exceptional V-final language (i.e. with the possessor pronoun following the noun) is again Oromo (see also sections 9.2.2.2 and 9.2.3):

Oromo (Stroomer 1987: 123)

(63) aabaa-ni keena gongooni taay-a	
father-S	our Gongoni live-3SG.M.PRES
'Our father lives at Gongoni'	

(p.306)

Table 9.7. Basic word order and the position of the possessor pronoun relative to the noun

Ng: 6 lgs. gN: 19 lgs. gN/Ng: 4 lgs.

VO: 14 lgs.	4 (+4 sx)	2 (+2 sx)	1 (+1 ax)
OV: 30 lgs.	1 (+6 sx)	17 (+5 px)	1
VO/OV: 5 lgs.	1	(2px)	2

Table 9.8. Possessor affixes and possessor NPs (G) in twenty languages

	GN	NG	GN/NG
g-N (px)	9		
N-g (sx)	4	5	1
g-N/N-g		1	

9.4.2.2. The relationship between possessor pronouns and possessor NPs

On the basis of an investigation of possessive constructions in seventy-five languages, Ultan formulated several general statements, such as (Ultan 1978c: 36):

Personal possessive prefixes always imply a GN order, but not the converse; for personal possessive suffixes there is no such rule.

This is generally confirmed by the languages in my sample (see Table 9.8). Although not all details are clear, Pipil (g-N/N-g and NG) could be a counterexample (cf. note 35 in Chapter 6). This language has prefixes (see example below) as well as two ‘possessive suffixes’, *-yu* and *-w(an)*, which occur with e.g. body-parts and kinship relations (Campbell 1985: 42–6). Note, however, that languages with both prefixes and suffixes are not logically excluded in Ultan’s formulation.

Pipil (Campbell 1985: 118)

(64)	i-kal	nu-te:ku
	his-house	my-father [i.e. g-N G[= g-N]]
‘my father’s house’		

I have no counter-examples to the rule which states (Ultan 1978b: 36):

GN constituent order in a nominally possessed construction implies the same order in a pronominal (non-affixal) construction.

Although in my sample there is perfect correlation between the position of nominal and (free) pronominal possessors, Ultan (1978c: 24) found that there are some languages with different orders. In these languages it is always the case that (p.307)

Table 9.9. Free possessor pronouns (g) and possessor NPs (G) in twenty-nine languages

	gN	Ng	gN/Ng
GN	19		

the pronominal form precedes, whereas the nominal form follows the noun (i.e. gN and NG).

Finally, Ultan (1978c: 36) also found that ‘a free, less intimate possessor form may imply a bound, more intimate form, but not the converse’, which is an instance of iconicity in language (Chapter 8; cf. Fox 1981) and relates to languages which have both free and bound attributive possessor pronouns and in which one set is used with ‘inherent possession’ and the other in the case of ‘non-inherent possession’. In my sample Burushaski, Gude, Ngiti, and Samoan satisfy both conditions and it appears that each of these languages confirms Ultan’s observation (section 6.3.1.1).

9.4.3. Relative clauses

For the relative clause the correlation is stronger in the group of VO-languages than in the OV-languages. Thirteen OV-languages have RelN order and eight have NRel order; in the group of VO-languages, on the other hand, eleven languages have NRel order, whereas only two have RelN order: Ngiti and Tsou.

There are two reasons that could explain the weak correlation between RelN and OV order. First, it is likely that ‘heaviness’ contributes to the number of OV-languages in which the relative clause follows rather than precedes the noun (section 9.3.3.4).³⁴ Secondly, it is perhaps not a coincidence that nearly all languages with correlative and adjoined relative clauses belong to the OV group (Bambara, Hittite, Ngalakan; Nunggubuyu has been classified as VO/OV/free in Chapter 8). In other words, apart from effects due to ‘heaviness’, the correlation between OV and RelN might have been stronger if the group of OV-languages had not also contained the languages with the non-integrated relative constructions.

9.4.4. Conclusion

The first hypothesis that was formulated on the basis of the *Principle of Head Proximity* (concerning adjective-noun adjacency) was generally confirmed. The (p.308) only problematic language was Oromo, which also shows some other unusual ordering properties (sections 9.4.2 and 9.4.2.1). It turned out that the second hypothesis, concerning the position of adnominal modifiers relative to the head noun, mostly holds for embedded noun modifiers G and Rel. Dryer’s data do not clearly indicate that the *Principle of Head Proximity* significantly influences the position of demonstratives, numerals, and adjectives relative to the noun. In the languages of my variety sample, however, demonstratives do tend to pattern according to the *Principle of Head Proximity*. Numerals on the other hand have a tendency to precede the noun and adjectives show no preference for pre- or postnominal position.³⁵

Considering all the word order phenomena (both inside and outside the NP) that support the idea that in a domain the head constituent prefers to occur as closely as possible to the head of its superordinate domain, we may conclude that the *Principle of Head Proximity* enables us to explain quite a variety of universal word order tendencies, but that the occurrence of non-embedded noun modifiers (dem, num, A)

between the head noun and the main verb does not seem to pose too many problems in terms of head proximity and ultimately for the human language processor (section 9.1). A possible explanation for this difference in syntactic behavior between embedded and non-embedded modifiers has already been given above (section 9.3.3.5), where it was also hypothesized that it is perhaps not so much the linguistic form that is important here (embedded vs. non-embedded modifier), but the fact that possessor NPs and relative clauses are complex, referential structures, which presumably are more difficult to process by the human parser than simple noun modifiers such as demonstratives, numerals, or adjectives.

(p.309) 9.5. The position of embedded modifiers

The *Principle of Head Proximity* says that the head of any domain tends to be contiguous with the head of the superordinate domain. In other words, strictly speaking it does not distinguish between the head of an embedded domain (possessor NP, relative clause) and the head of a non-embedded domain (adjective). The fact that one usually finds the adjective next to the noun, and not (the head of) some other subordinate domain in the NP (should they both precede or follow the noun), can also be attributed to Head Proximity, since this results in better overall head proximity ratios (see section 9.3.4.2.2). Furthermore, the *Principle of Domain Integrity* (Chapter 8) states that embedded domains must not occur in between constituents of the matrix domain (the *Principle of Scope* also plays a role here; see section 10.3). The *Principle of Domain Integrity* also accounts for the fact that the head of the embedded modifier is not expressed discontinuously from its dependants, so that it alone would be able to comply with the *Principle of Head Proximity*.

These are the preferred patterns according to the *Principle of Head Proximity* and the *Principle of Domain Integrity* in a head-initial and in a head-final language:

(65) V[N A [V/N_{Head} ...]Embedded modifier]NP ... (head-initial)

(66) ...[[... V/N_{Head}]Embedded_modifier A N]NP V (head-final)

Observe that the *Principle of Head Proximity* also makes predictions about the linear organization of constituents of the embedded modifier: dependants in the embedded modifier should not occur in between the head of their own domain and the head of the superordinate domain N.

Although there is very little information about the position of embedded modifiers relative to the non-embedded noun modifiers in the matrix NP, it seems that there is at least one other factor involved concerning the position of embedded domains in the NP, namely heaviness. I only have the relevant data for a representative sample of the European languages and here we find interesting differences between the syntax of prenominal and postnominal noun modifiers: whereas embedded modifiers can be found in various positions in the NP when they precede the noun, they have a strong tendency to appear after all non-embedded modifiers when they follow the noun. In view of the fact that (especially in spoken language) one seldom uses more than two or three modifiers in the same NP (section 1.5.4.3), one can of course hardly speak of a basic or unmarked order in the case of such unusually complex NPs. Therefore I have restricted myself to what is deemed to be the least marked position of the possessor NP (G) and the relative clause (Rel) relative to constituents of the matrix NP (see Rijkhoff 1997: 361–2 for details about alternative ordering possibilities).

(p.310) (67) Position of prenominal embedded modifiers G and Rel in a representative sample of European languages:

G/Rel: Bashkir (dem num A N), Chuvash (dem num A N), Kalmyk (dem num A N), Nogai
 NP- (dem num A N), Turkish (dem num A N), Laz (dem num A N), Lezgian (dem num A
 initial N),³⁶ Basque (num A N dem)

G: NP- East Armenian (dem num A N), Finnish (dem num A N), Hungarian (dem num A N),
 initial Udmurt (dem num A N), Ossetic (dem num A N), Georgian (dem num A N)

dem	Abkhaz (dem num-N-A), Adyghe (dem N-A num), Nenets (dem num A N), Swedish [G]... N (dem num A N)
[Rel]...	
N	
dem	Ingush (possibly also: A [G] N; Johanna Nichols, personal communication)
num	
[G] A N	
dem	Lithuanian (dem num A N)
num A	
[G] N	
dem	Nenets (dem num A N)
num A	
[Rel] N	

(68) Position of postnominal embedded modifiers G and Rel in a representative sample of European languages:

G/Rel:	Assyrian (dem num N A), Maltese (dem num N A), Dargwa (dem num A N)?, ³⁷ Svan
NP-final	Bulgarian (dem num A N), Dutch (dem num A N), English (dem num A N), Greek (dem num A N), Icelandic (dem num A N), Polish (dem num A N), Russian (dem num A N), Albanian (dem num N A), Kirmanji (dem num N A), Vlax Romani (dem num A N), Romanian (dem num N A), Sardinian (dem num N A), Spanish (dem num N A), Breton (num N A dem), Irish (num N A dem), Tati (dem num A N?)
Rel:	East Armenian (dem num A N), Hungarian (dem num A N), Finnish (dem num A N),
NP-final	Udmurt (dem num A N), ³⁸ Georgian (dem num A N), Swedish (dem num A N), Lithuanian (dem num A N)

(p.311) These data show that in most European languages the embedded modifier does not appear in between constituents of the matrix NP and that in all languages with more than one postnominal modifier (in italic print) the embedded domain always appears after all other modifiers, which is fully in accordance with the *Principle of Domain Integrity* (see also Lehmann 1984: 201). Violations of the *Principle of Domain Integrity* are all restricted to embedded domains in prenominal position. Taking into account that most European languages have the demonstrative, numeral, and adjective before the noun, this asymmetry is probably due to heaviness (section 9.3.3.4). In the case of a noun-final structure the tendency for embedded (i.e. heavy or complex) modifiers to appear after all other modifiers only results in violations of the *Principle of Domain Integrity*. But once the embedded modifier occurs in postnominal position, its tendency to follow all (less heavy) non-embedded modifiers is in perfect agreement with the *Principle of Domain Integrity*.

Appendix

According to the PHP all head nouns in a clause prefer to be contiguous with the main predicate, whereas of course in the actual utterance at best two nouns can occur in such a position. In other words, the greater the number of NPs in a clause (i.e. the greater the number of head nouns that will have to appear at some distance from the main predicate), the worse this is in terms of Head Proximity. On earlier occasions (Rijkhoff 1984, 1986) I have discussed a variety of phenomena which all seem to point to a general tendency (at least in the spoken language) to improve the overall ratio of head proximity either by reducing the number of NPs, or by increasing the number of verbs in a sentence; some examples of this tendency are given below. Whether or not this tendency is actually due to an effort to facilitate language processing remains to be tested, but it is perhaps significant that the opposite tendency, i.e. to try to include as many subordinate domains as possible in a superordinate domain, is definitely the marked alternative.

Note that the seriousness of problems in terms of Head Proximity at least partly depends on the parts-of-speech system and the morphological type of a language. Thus in ‘verby’ languages such as Cayuga (section 1.5.2) one expects there to be few problems in terms of head proximity, since ‘normal discourse can consist largely of verbs’ (Mithun 1992: 57).

As regards the relationship between head proximity and morphological type, it was mentioned in Chapter 8 (section 8.2.1) that in many languages it is unusual or sometimes nearly impossible to have two full lexical NPs in the same sentence and this holds in particular for languages of the polysynthetic type. In these languages the main predicate (p.312) often contains the (nominal and pronominal) arguments, witness:

Southern Tiwa (Whaley 1997: 131; original example from Allen et al. 1990):

(1) Ti-khwian-mu-ban

IS-dog-see-PAST

‘I saw the dog’

Another ‘strategy’ to improve head proximity is to increase the number of verbs (rather than reduce the number of NPs). This seems to be the case in languages that tend to distribute NPs over two or more clauses (as was already mentioned in section 1.5.4.3; see also Rijkhoff 1986: 108 f.). Consider the following remarks on Kobon, Usan, and Hix-karyana:

Kobon (Davies 1981: 45–6)

‘All of the obligatory and optional arguments can co-occur but if many arguments are involved there is a strong preference for dividing them over two or more clauses which may contain an identical predicate rather than including them all in one clause.’

Usan (Reesink 1984)

‘as a good Papuan language, [Usan] is characterized by a predominance of predicates over terms. ... clauses with more than two or three terms are very rare in Usan’ (*ibid.* 152).

‘most states of affairs are expressed by combinations of verbs, many of which are highly restrictive with regards to their semantic loading. This makes explicit reference to semantic functions in most cases redundant’ (*ibid.* 108).

Frequently, because of the restriction on number of terms per predicate, Instrument function] is expressed as Go[al function] of a preceding predication and understood in the predication that follows, as in:

Usan (Reesink 1984: 133–4)

(2) wo bei ba nam su-erei (=sorei)

he axe take tree cut-3SG.RemP

‘He took an axe and cut the tree’

Although this example comes very close to a serial construction, Reesink states it consists of two clauses. Finally, in Hixkaryana (Derbyshire 1979: 39): ‘there is a preference to limit the adjuncts in a single sentence to one and add further sentences, as needed, with a repetition of verbs.’

Notes:

- (1) When I proposed the *Principle of Head Proximity* (Rijkhoff 1984) I was unaware of the fact that a few years earlier Lyn Frazier (Frazier 1979; see also Frazier 1985: 146) had formulated an ordering principle with a rather similar name to account for some of the Greenbergian correlations:

Head Adjacency Constraint: In the unmarked case, major (recursive) phrases must follow the head of a mother phrase in prepositional languages and must precede the head in postpositional languages.

Although the notion ‘head’ is given a different interpretation (like Dryer and Hawkins—see below—Frazier assumes some version of a Chomskyan constituent structure tree) and although the two principles do not make the same predictions, it seems that she had essentially the same insight.

(2) I deliberately avoid ‘verb phrase’ here, because this term has been more or less annexed by Chomskyan-inspired grammatical theories, where it refers to structures that are less than a full clause.

(3) For example, the strong version of the *Principle of Head Proximity* also accounts for the fact that auxiliary verbs appear before the main verb in V-initial languages, after the main verb in V-final languages, and in clause-second position or after the main verb in non-rigid V-final languages (Rijkhoff 1984, 1986, 1987; cf. Dik 1997: 413–14); see also below on Dryer’s word order correlations.

(4) See also note 1 and Dryer’s processing account of his Branching Direction Theory (Dryer 1992: 128).

(5) Compare also English ‘I have never seen so beautiful a girl.’

(6) See also Hetzron (1978b: 171–2) on ‘big’-fronting. He found that in a number of languages (e.g. Amharic, Japanese, Spanish, Romanian) there is a tendency to place the size adjective first in the line of adjectives. ‘The likely explanation of this fronting of the adjective “big” is that size-qualifications are obligatorily treated as emphatic in these languages.’

(7) Regarding Turkish *bir*, Lewis (1967: 54) writes: ‘When it serves as an indefinite article, *bir* usually comes between adjective and noun: *büyük bir tarla* “a large field”, *güzel bir bahçe* “a beautiful garden”. When it means “one”, it must precede the adjective, just like any other numeral: *iki küçük tarla satti, bir büyük tarla aldi* “he sold two small fields, he bought one large field”. This must not be taken to imply that *bir* when it precedes an adjective and noun is always translated by “one”; English idiom may sometimes call for “a” or “any” ... The key to understanding this point lies in the basic principle of Turkish syntax: whatever precedes, qualifies. The essential difference between *güzel bir bahçe* and *bir güzel bahçe*, both of which translate “a beautiful garden”, is that the first means a beautiful member of the class “garden”, the second a member of the class “beautiful garden”, *güzel bir bahçe* is a beautiful garden as distinct from a less beautiful or even frankly ugly garden; *bir güzel bahçe* is a beautiful garden as distinct from a beautiful meadow or an ugly forest.’ See also van Schaik (1996) on the internal syntax of Turkish NPs.

(8) The various forms for ‘one’ in Nasioi are probably derived from demonstrative forms, because the former consist of initial *n-* plus the appropriate demonstrative form; thus *náve* [one:C15] ‘one (tree)’ vs. *ave* [this:C15] ‘this (tree)’. The ‘absolute form’ (class 1) is used in counting and in indefinite NPs.

(9) There is no formal distinction between attributive and predicative adjectives (see also example (12) above) and since copulas are absent in Nasioi the adjective in at least some of these examples can also be regarded as the main predicate, yielding ‘this tree is bad’ (Rausch 1912: 132).

(10) Cf. Senft (1986: 43) on Kilivila, where demonstratives, numerals, and (certain) adjectives ‘require concord with the class of the noun they refer to. This concord is secured by a special class of formatives, the “nominal classifiers” or “classificatory particles”, that represent the system of noun classification.’ Consider, for example:

Kilivila (Senft 1986: 69)

(1)	mi-na-si-na	na-yu	na-manabweta	vivila
	this-F-Pl-this	F-two	F-beautiful	girls
‘these two beautiful girls’				

Once the head noun has been mentioned these modifiers can also occur alone. Then the class marker on the modifier serves as an anaphoric element and secures discourse cohesion (Senft 1986: 106; see also Senft 1985).

(11) There is no formal distinction between attributive and independent (pronominal) demonstratives in Oromo.

(12) See also Tucker (1994: 245) on so-called ‘adjectives’ requiring the relative particle *ma-* in Nilo-Saharan Kenya Luo (or Dholuo). On the absence of a true class of adjectives in Bantu languages, see e.g. Welmers (1973); Hagège (1974); and Bot Ba Njock (1977).

(13) See also Creider and Creider (1989: 64) on numerals in Nandi (Nilo-Saharan): ‘Numbers are morphologically nouns.’

(14) The text on noun modifiers in Beja continues as follows (Hudson 1976: 106): ‘However, there are two other differences between preceding and following modifiers: if the modifier precedes, the noun must not contain the definite article (unless the modifier is one of the demonstratives ...; and the modifiers have different concordial suffixes according to whether they precede or follow.’

(15) On asymmetry in linguistics, see also Tesnière (1959: 22, 32 f.); Yngve (1960: 465); Schwartz (1971); and Lehmann (1982b: 263).

(16) Ultimately Hawkins (1983: 75, 86) was able to collapse various implications into two statistical implications with an adposition (preposition, postposition) as ultimate antecedent:

Prepositional Noun Modifier Hierarchy:

Prep ⊃ ((NDem V NNum ⊃ NA) & (NA ⊃ NG) & (NG ⊃ NRel))

Postpositional Noun Modifier Hierarchy:

Postp ⊃ ((AN V RelN ⊃ DemN & NumN) & (DemN V NumN ⊃ GN)).

(17) Tigre (Tosco 1998) belongs to Hawkins’s type 18 (SOV & Pr & NG & NA), of which Hawkins (1983: 285) had no examples; the same goes for Tsou (Szakos 1994), which belongs to Hawkins’s non-attested type 4, i.e. VI & Pr & GN & NA (Hawkins 1983: 283).

(18) To be fair, Dryer makes it clear at various places that he is not unaware of this problem himself (cf. Dryer 1992: 84 n. 2, 96 n. 12, 120; also Dryer 1989b: 870).

(19) The revised and alternate versions of the Branching Direction Principle read as follows (Dryer 1992: 114, 116):

BRANCHING DIRECTION THEORY (revised version): Verb patters are non-phrasal categories or phrasal categories that are not fully recursive, and object patters are fully recursive phrasal categories in the major constituent tree. That is, a pair of elements X and Y will employ the order XY significantly more often among VO languages than among OV languages if and only if X is not a fully recursive phrasal category in the major constituent tree and Y is a fully recursive phrasal category in the major constituent tree.

BRANCHING DIRECTION THEORY (alternate version): Verb patters are heads and object patters are fully recursive phrasal dependents. I.e., a pair of elements X and Y will employ the order XY significantly more often among VO languages than among OV languages if and only if X is a head and Y is a phrasal dependent of X.

(20) Recall that in FG the functions Subject and Object are not universal, in that they are not necessarily part of the grammar of every language. Dik (1997: 417–19) has argued that these syntactic functions are probably irrelevant in languages that are deemed to have OS order.

(21) The difference between lexical and grammatical elements is considered rather fundamental in FG (Dik 1997: 159), and is reflected (among other things) in the expression component of the FG model: some ordering principles apply to lexical items (e.g. the *Principle of Head Proximity*), whereas others only relate to grammatical elements (Dik 1997: 413).

(22) Fijian has two articles: *a* (with common nouns) and *o* (with proper nouns).

(23) Hawaiian seems to be an example of a language that has the same group of elements figuring in two correlation pairs, since Dryer (1989a: 872) writes that this language belongs to the kind in which ‘plural

words are articles by virtue of the fact that they simultaneously code notions like definiteness or noun class and exhibit properties of articles as noun-phrase markers described above ... The plural word *nā* is an article, as illustrated by the fact that it occurs in article position, and by the fact that it codes definiteness: *nā* is the plural of the singular definite article *ka/ke*, as illustrated by *ke keiki* ‘the child’, and *nā keiki* ‘the children’ (Hawkins 1982: 8).²⁴

(24) Recall that possessor NPs and relative clauses typically provide the entity through which the referent of the matrix NP can be identified (the Identifying Entity or IE; see section 6.1).

(25) The expanded version of the principle of *Early Immediate Constituents* reads as follows (Hawkins 1994: 78–9): ‘The human parser prefers linear orders that maximize the IC-to-non-IC ratios of constituent recognition domains. Orders with the most optimal ratios will be preferred over their non-optimal counterparts in the unmarked case; orders with non-optimal ratios will be more or equally preferred in direct proportion to the magnitude of their ratios. For finer discriminations, IC-to-non-IC ratios can be measured left-to-right.’

(26) In his monograph, Hawkins (1994: 120) ascribes the *Principle of Head Proximity* and the *Principle of Domain Integrity* to Simon Dik, despite explicit statements (Dik 1997: 402) that they were first formulated by the present author. This is corrected in Hawkins (1997: 750).

(27) Only in case of the *Relator Principle* does Hawkins admit that theoretical postulates from Generative Grammar are relevant in this comparison, as when he writes that he works ‘on the assumption that Relators are constructing categories in our sense’ (Hawkins 1994: 120), which obviously they are not in Functional Grammar.

(28) On the interaction between EIC and other ordering principles, see also Hawkins (1997:759): ‘EIC is not the only processing principle regulating linear ordering proposed in PTOC [=Hawkins 1994—JR], though it is the major one. In addition, the principles of Promotion Attachment and Promotion Construction predict the possibility of certain focus positions occurring immediately adjacent to a verb. And the principle of Immediate Matrix Disambiguation predicts, in conjunction with EIC, a number of left-right asymmetries involving the positioning of sentential nodes across languages, e.g. the rightward skewing in favour of postnominal versus prenominal relative clauses.’

(29) Notice that Hawkins only deals with a subset of Dryer’s correlation pairs, whereas I have discussed all of them (see section 9.3.2).

(30) See Rijkhoff (1997; also section 9.5) on the position of the relative clause relative to the adjective and other modifier categories in the NP in the languages of Europe. Not surprisingly, I do not agree with Hawkins’s statement (1994: 269) that ‘adjectives and relative clauses perform a very similar semantic function as (restrictive or appositive) modifiers of the head noun’. I have argued in Chapters 4 and 6 that adjectives typically serve as *qualifying satellites*, whereas relative clauses (notably in those languages that actually have a class of adjectives) commonly serve as *location satellites*.

(31) I will also ignore the position of the category that ‘constructs’ the relative clause, as discussed in Hawkins (1994: 270), since this is another feature of the Chomskyan approach to grammatical theory that does not have a counterpart in other theories.

(32) In Table 9.5 only Tsou adjectives are bracketed, but Stan Starosta (personal communication) has suggested that other modifier categories are not part of a simple, integral NP either. He paraphrased the Tsou equivalent of the English NP ‘these two big dogs’ as follows: ‘the two ones there which are big ones which are dogs.’

(33) Whether or not a language permits doubling of the adjective seems to be partly determined by the parts-of-speech system of that language (Hengeveld et al. forthcoming).

(34) On the role of complexity/heaviness in syntax, see e.g. Mallinson and Blake (1981: 157); Hawkins (1983: 89); Siewierska (1988: 43–7); Dik (1997: 430–1).

(35) The fact that cross-linguistically adjectives do not seem to have a preferred position relative to the noun may have to do with Clark and Clark’s second strategy (mentioned earlier in sections 9.1 and 9.3.4.2, and repeated here for convenience):

STRATEGY 2: After identifying the beginning of a constituent look for content words appropriate to that type of constituent.

A simple or embedded NP cannot contain more than two kinds of content words, a noun and possibly an adjective, and since normally an adjective immediately precedes or follows the noun (see section 9.2), it does not matter a whole lot for processing purposes which content word is encountered first. If the first content word is a noun, the processor encounters the content word that is most appropriate to the NP. If the adjective is encountered first, the processor ‘knows’ that the next content word will be the noun. Thus, for processing purposes it does not really matter whether the adjective precedes or follows the noun: in either case the listener will be able to successfully apply Clark and Clark’s Strategy 2. Support for this theory can also be found in Hengeveld et al. (forthcoming), in which it is argued that the order of noun-adjective is highly relevant for processing purposes in those languages where adjectives cannot be distinguished from nouns and/or verbs. This is confirmed here: Samoan, Hurrian, Imbabura Quechua, and Turkish all have flexible adjectives, which do not occur in between the head of the NP and the main predicate.

(36) G and Rel tend to precede all other modifiers in Lezgian, but a short (and ‘non-referential’) genitive and relative clause can follow other modifiers (Haspelmath 1993: 260–2); see also Kazenin (1994: 148) on the position of G and Rel in Godoberi).

(37) In Dargwa G and Rel do not occur on the same side of the noun together.

(38) Udmurt (like other Finno-Ugric languages such as Finnish and Hungarian) has a finite relative construction that follows the noun and a participial construction that must precede the noun. The participial construction has no fixed position relative to the other prenominal modifiers. According to Maria Vilkuna (personal communication) ‘the participial construction can be seen as an adjective phrase, but earlier placement is preferred than for plain adjectives, at least when it comes to acceptability judgments, to avoid ambiguities’.

10 The Principle of Scope

10.1. Introduction

The *Principle of Scope* states that the semantic distance of grammatical and lexical modifiers (or, in FG, *operators* and *satellites*, respectively) relative to the head in the underlying structure is iconically reflected in the actual linguistic expression:

The Principle of Scope

Modifiers tend to occur next to the part of the expression that they have in their scope.

The *Principle of Scope* accounts for two ordering features. First, it predicts—implicitly—that constituents that are in the scope of a certain modifier (i.e. elements that are part of the same layer in the underlying NP structure) are expressed in a continuous sequence. Secondly, it predicts—explicitly—that operators and satellites occur immediately before or after the material they have in their scope. Recall that, despite the format of the representation, an underlying structure has no *linear* order; what counts is the hierarchical organization of the layers. In the present context a simplified version of the underlying structure will suffice (for the fully specified underlying structure of the NP I refer the reader to Chapter 7):

(1) $[\omega_{2b} [\omega_{2a} [\omega_1 N \tau_1] \tau_{2a}] \tau_{2b}]$

In (1) $[\omega_1 N \tau_1]$ constitutes the quality layer, containing the qualifying operator ω_1 and qualifying satellite τ_1 (Chapter 4), as well as the element they have in their scope, the head noun N (Chapter 2). The quality layer is nested in the quantity layer with quantifying operator ω_{2a} and quantifying satellite τ_{2a} (Chapter 5). In its turn the quantity layer is nested in the location layer with localizing operator ω_{2b} and localizing satellite τ_{2b} (Chapter 6). In the unrealistic case that an NP might be specified for all its operator and satellite positions, these would be some of the patterns in accordance with the *Principle of Scope* (but recall that, for example, numerals (ω_{2a}) and nominal aspect markers (ω_1) are normally mutually exclusive; section 4.2.1.2):

(2) $\omega_{2b} \omega_{2a} \omega_1 N \tau_1 \tau_{2a} \tau_{2b}$

(3) $\tau_{2b} \tau_{2a} \tau_1 N \omega_1 \omega_{2a} \omega_{2b}$ (the mirror pattern of (2))

(4) $\omega_{2b} \tau_{2a} \omega_1 N \tau_1 \omega_{2a} \tau_{2b}$

(5) $\tau_{2b} \omega_{2a} \tau_1 N \omega_1 \tau_{2a} \omega_{2b}$ (the mirror pattern of (4))

(6) $\tau_{2b} \omega_{2b} \tau_{2a} \omega_{2a} \tau_1 \omega_1 N$

(p.314)

(7) $N \omega_1 \tau_1 \omega_{2a} \tau_{2a} \omega_{2b} \tau_{2b}$ (the mirror pattern of (6))

(8) $\omega_{2b} \tau_{2b} \omega_{2a} \tau_{2a} \omega_1 \tau_1 N$

(9) $N \tau_1 \omega_1 \tau_{2a} \omega_{2a} \tau_{2b} \omega_{2b}$ (the mirror pattern of (8))

The *Principle of Scope* does not permit us to make any predictions about the relative order of same-level operators (ω) and satellites (τ) if they both precede or follow the head noun.¹ Thus, the ordering patterns (6)-(9) are all equally acceptable in terms of the *Principle of Scope*. According to the *Principle of Domain Integrity*, however, all embedded domains (such as possessor NPs and relative clauses, symbolized by τ_{2b}) must occur in the periphery of the NP, i.e. before or after all other constituents of the matrix NP. Interactions between the *Principle of Scope* and the *Principle of Domain Integrity* affecting the syntax of embedded noun modifiers will be discussed in section 10.3.² First, however, I will investigate word order patterns in the simple NP in the light of the *Principle of Scope*.

10.2. The relative order of modifiers in non-complex NPs

A simplified representation of the underlying structure of a simple, integral NP looks like this:

(10) $[\omega_{2b} [\omega_{2a} [\omega_1 N \tau_1]]]$

The qualifying satellite in a simple NP (τ_1) can only be a non-embedded subordinate domain, i.e. an adjective; another specification of the qualifying satellite, such as a relative clause (i.e. an embedded subordinate domain), would result in a complex NP.

The *Principle of Scope* only permits the following subset of all logically possible orderings (where all modifiers are free, integrated constituents of the NP):

(11)

- a. $\omega_{2b} \omega_{2a} \omega_1 N \tau_1$
- b. $\omega_{2b} \omega_{2a} \tau_1 N \omega_1$

(12)

- a. $\tau_1 N \omega_1 \omega_{2a} \omega_{2b}$ (the mirror patterns of (11a–b) respectively)
- b. $\omega_1 N \tau_1 \omega_{2a} \omega_{2b}$

(13)

- a. $\omega_{2b} \omega_1 N \tau_1 \omega_{2a}$
- b. $\omega_{2b} \tau_1 N \omega_1 \omega_{2a}$

(14)

- a. $\omega_{2a} \tau_1 N \omega_1 \omega_{2b}$ (the mirror pattern of (13a–b) respectively)
- b. $\omega_{2a} \omega_1 N \tau_1 \omega_{2b}$

(p.315)

(15)

- a. $\omega_{2b} \omega_{2a} \tau_1 \omega_1 N$

b. $\omega_{2b} \omega_{2a} \omega_1 \tau_1 N$

(16)

- a. $N \omega_1 \tau_1 \omega_{2a} \omega_{2b}$ (the mirror pattern of (15a–b) respectively)
- b. $N \tau_1 \omega_1 \omega_{2a} \omega_{2b}$

We can make these patterns more transparent by substituting the symbols by the appropriate modifier categories demonstrative (ω_{2b}), numeral (ω_{2a}), adjective (τ_1), and free nominal aspect marker (= asp; ω_1)

(17) a. dem num asp N A

b. dem num A N asp

(18) a. A N asp num dem

b. asp N A num dem

(19) a. dem asp N A num

b. dem A N asp num

(20) a. num A N asp dem

b. num asp N A dem

(21) a. dem num A asp N

b. dem num asp A N

(22) a. N asp A num dem

b. N A asp num dem

As already stated in section 2.2.2, grammatical expressions of numerals and overt expressions of nominal aspect operators are usually mutually exclusive, so that the following NP patterns only contain four constituents. These are the possible patterns according to the *Principle of Scope* with a free nominal aspect operator and without a numeral.

(23) dem asp A N dem A N asp asp A N dem A N asp dem

dem asp N A dem N A asp asp N A dem N A asp dem

These are the patterns with a numeral and without a nominal aspect marker:

-
- (24) dem num A N dem A N num num A N dem A N num dem
 dem num N A dem N A num num N A dem N A num dem

Below I will first investigate the syntax of constituents that make up the quality layer on the underlying structure of the NP (noun, adjective, nominal aspect marker), which according to the *Principle of Scope* should always occur in an uninterrupted sequence (see above). Although the position of adjectives has already been discussed in Chapter 9, I will briefly return to this in sections 10.2.1 and 10.2.2. The syntax of adnominal demonstratives and numerals is discussed in section 10.2.3.

(p.316) 10.2.1. Free nominal aspect markers and adjectives

In section 4.2.1.2 it was established that in the great majority of languages the nominal aspect operator (ω_1) is realized as an inflectional category on the noun, which means that its expression is determined by morphological rather than syntactic rules. There are, however, seven languages in the sample with what may be regarded as a free nominal aspect marker: Galela, Guaraní, Hixkaryana, Hmong Njua, Ika, Koasati, and Sarcee. For these languages the *Principle of Scope* predicts that the nominal aspect marker (if part of the integral NP) immediately precedes or follows the noun.

-
- (25) a. ω_1 N (or asp N)
 b. N ω_1 (or N asp)

Should, however, the NP also contain an adjective which appears on the same side of the noun as the free nominal aspect marker, then either order is permitted according to the *Principle of Scope*:

-
- (26) a. $\tau_1 \omega_1$ N (or A asp N) or $\omega_1 \tau_1$ N (or asp A N)
 b. $\omega_1 \tau_1$ N (or asp A N) or N $\tau_1 \omega_1$ (or N A asp)

That is to say: in the simple, integral NP a free nominal aspect may only be separated from the noun by an adjective (see below on the impact of the *Principle of Head Proximity* in such cases). I will first investigate the position of the free nominal aspect marker in languages without adjectives and then turn to their position in languages that do have a class of adjectives.

10.2.1.1. The position of aspect markers in languages without adjectives

At least four of the seven languages mentioned above in section 10.2.1 lack a class of (free) adjectives: Galela, Hixkaryana, Koasati, and possibly Sarcee.³ If a language has no adjectives one may expect the nominal aspect marker always to occur immediately before or after the noun.

10.2.1.1.1. Galela

Apparently no other modifier can occur in between the so-called plural marker *bi* ('collective' in my interpretation) and the head noun. *Bi*, which only appears

in definite NPs, is not just mutually exclusive with numerals, but also with demonstratives and possessor pronouns (the two latter categories also precede the noun):

Galela (van Baarda 1908: 32)

(27) o	bi	dòro
ART	BI	garden
'gardens'		

(p.317) 10.2.1.1.2. Hixkaryana

Hixkaryana *komo* ‘collective’ follows the noun but like all other ‘modifying particles’ (Derbyshire 1979: 45) that follow the noun it has no fixed position. The grammatical status of *komo* is not clear and it may be interesting to note that *komo* and most of the other modifying particles can also occur by themselves, but ‘in these cases they always refer to some noun in the context and are considered to be modifying an ellipsed nominal head’ (Derbyshire 1979: 83). Note furthermore that elements such as *komo* ‘collective’, *heno* ‘now dead, the late ...’, and *rma* ‘same referent’ have properties of clitics as well as free elements.⁴ Phonologically they are bound forms, but morpho-phonologically and distributionally they behave like free constituents.

Hixkaryana (Derbyshire 1979: 83)

(28) n-amryek-yatxkon	heno	komo	rma	hati,	kamarayana	komo
3S-hunt-DistPC.COLL	dead	COLL	SR	HSY	jaguar_person	COLL
[Apparently] those jaguar people, now dead, used to still go hunting'						

Alternative orderings of the constituents in the string before the tail constituent *kamrayana komo* are:

Hixkaryana (Derbyshire 1979: 83)

(29) namryekyatxkon	kom	heno	rma	hati,	...
COLL	dead	SR	HSY		
(30) namryekyatxkon	heno	rma	kom	hati,	...
dead	COLL	SR	HSY		
(31) namryekyatxkon	komo	rma	heno	hati,	...
dead	COLL	SR	HSY		

Due to lack of detailed information it is not possible to determine whether or not Hixkaryana constitutes a true counter-example to the prediction based on the *Principle of Scope*, but in view of Derbyshire’s observations it seems at least possible that *komo* is not an entirely free and/or fully integrated noun modifier.

10.2.1.3. Koasati

A suffix expresses both the diminutive and singulative aspect in Koasati, but with non-human animate nouns the free form *lawista* ‘small’ is used to express both the diminutive and ‘plural’ (= collective aspect in my interpretation; see also section 4.2.1.3).

Koasati (Kimball 1991: 460)

(32) nita-sí

bear-DIM

‘bear cub’

(p.318)

(33) nitá lawísta

bear small[PL]

‘bear cubs’, lit. ‘little bears’

Presumably *lawísta* normally appears immediately following the noun, since Kimball (1985: 394) writes that it frequently fuses with the preceding noun.

10.2.1.4. Sarcee

In Sarcee nouns may occur with *-ká* or *yìná*, which indicate ‘plurality or collectivity’ (Cook 1984: 65), as well as deference. The use of the suffix *-ká* is mostly restricted to human nouns (particularly kinship nouns). Interestingly the suffix *-ká* may also appear on a numeral or a relative clause ‘as a kind of agreement’ (Cook 1984: 65–6), but only if the head of the NP is affixed by *-ká* as well, for example:

Sarcee (Cook 1984: 65)

(34) gūnìsnóní-ká tñí-ká nàníló-là

nine-KÁ dog-KÁ she-laid

‘She gave birth to nine dogs’

The other form, *yìná*, with which we are concerned here, is employed with other nouns, such as personal names, but also with demonstratives (e.g. *dìná yìná* ‘these people’). As to its grammatical status, Cook (1964: 65) writes that *yìná* ‘never occurs alone as a noun, but it is not an affix either’; I have found no examples with *yìná* separated from the noun.

10.2.1.2. The position of aspect markers in languages with adjectives

The following languages in the sample have free nominal aspect markers as well as a class of adjectives: Guaraní, Hmong Njua, Ika, and possibly Turkish.

10.2.1.2.1. Guaraní

Guaraní has two elements that may be regarded as collective aspect markers; both are characterized as ‘qualifiers’, a wastebasket category containing various elements, most of which have no fixed position in the NP. One of the collective aspect markers may also be used with verbs; the other can occur with nouns as well as pronouns. As a rule they occur in NP-final position, after all other ‘modifiers’, such as *té* ‘superlative’, *nte* ‘merely, just’ (which may also occur as a clausal constituent), and the postpositional phrase (Gregores and Suárez 1967: 144, 150). Normally the noun does not occur with more than three dependants (if they are that) in Guaraní and only the prenominal modifiers have a fixed position (illustrating again the asymmetry between prenominal and postnominal forms of modification; see section 9.2.3):

Guaraní (Gregores and Suárez 1967: 150)

(35) upé la ?óga nte ità gwigwà kwéra
that the house just stone of PL
‘just those houses of stone’

(p.319) Since there is insufficient information about the morpho-syntactic status of *kwéra*, it is not possible to say whether or not Guaraní constitutes a counterexample to my predictions. However, the data presented by Gregores and Suárez (1967: 144, 150) suggest *kwéra* might be a (noun) phrase final clitic or an appositional element rather than a fully integrated noun modifier.

10.2.1.2.2. Ika

As in the other South American languages Guaraní and Hixkaryana, the element in question (here *džina*—glossed as ‘plural’ in Frank’s description of the language, but see section 4.2.1.2) appears after the noun. It is categorized as a quantifier and since there are no details about its morpho-syntactic status and position relative to other postnominal modifiers (numerals, adjectives, modifying NPs) it is not possible to say whether *džina* always conforms to the preferred ordering in terms of the *Principle of Scope* (recall that most adnominal ‘adjectives’ occur with *kawa* ‘seem’; Frank 1990: 32).

Ika (Frank 1990: 29)

(36) nΛ-gunamí džina
1-worker PL
‘my workers’

10.2.1.2.3. Hmong Njua

Hmong Njua *cov* ‘collective’ has already been discussed in section 4.2.1.2. Other prenominal modifiers are possessor NPs and pronouns, as well as numerals of course, but apparently *cov* immediately precedes the noun:

Hmong Njua (Harriehausen 1990: 115)

(37) kuv cov hoob

1SG COV room

'my rooms'

10.2.1.2.4. Turkish (or *bir* revisited)

In section 9.2.1.3 the position of the Turkish so-called indefinite article *bir* between the adjective and the noun was attributed to emphasis on the adjective. A recent study by Schroeder suggests, however, that there is an alternative explanation for the position of 'indefinite' *bir* in the NP. Schroeder (1999) investigated Turkish noun phrases in spoken discourse and found that indefinite *bir* has various functions, some of which depend on the kind of noun with which it combines; here I will restrict myself to *bir* in combination with nouns that are used to refer to discrete spatial entities. Apparently in such cases *bir* is 'not to be primarily understood as an indefiniteness marker' (Schroeder 1999: 75), instead it mainly seems to serve two (p.320) other purposes (Schroeder 1999: 55–70):⁵

1. to introduce a new referent into the discourse (notably one that will be referred to again by anaphoric reference)—hence the association with the notion indefiniteness;
2. to mark the referent as an individual (recall that Turkish nouns are transnumeral).

Turkish (Schroeder 1999: 62)

(38) Abi-m	de	bir	ut	al-dl	...
	brother-POS.1SG	and	BIR	lute	get-PAST.1
'And my brother got a lute ... from Arabia					
Arabistan-dan ...	O	ut-u		çal-ryor	...
Arabia-ABL		that	lute-ACC	play-PRES	...
'He plays that lute					

Turkish (Schroeder 1999: 68)

(39) bir adam gel-iyor.	Onun	ad-1	da	Yahya	Kemal
BIR man come-PRES.	his	name	and	Yahya	Kemal
'There comes a man. And his name is Yahya Kemal'					

Schroeder's investigation indicates that indefinite *bir* is perhaps better analyzed as a singulative aspect marker (since 'it restricts the set of the referent to the individual'—Schroeder 1999: 62), whose occurrence is determined by pragmatic conditions (i.e. the introduction of new referents).

It is also true, however, that occasionally *bir* appears in NPs that do not introduce new referents (Schroeder 1999: 75, 85). Schroeder discusses three such cases, only one of which is relevant in the present context: the use of indefinite *bir* in combination with

modified nouns. He first establishes that *bir* is used in predicative NPs indicating class membership:

Turkish (Schroeder 1999: 92)

(40) Nazim Hikmet bir şair

Nazim Hikmet BIR poet

'Nazim Hikmet is a poet'

Apparently, *bir* still has an individualizing function in a predicate NP, because Schroeder (*ibid.*) states that here 'the article indicates that the referent of the phrase is *an individual member... of the class* of the referent'. Subsequently he discusses modified nouns and claims that indefinite *bir* is never grammatically required in such cases (although it may be quite appropriate in certain (p.321) contexts):

Turkish (Schroeder 1999: 92)

(41) Nazim Hikmet meşhur bir şair

Nazim Hikmet famous BIR poet

'Nazim Hikmet is a famous poet'

The appearance of *bir* with modified nouns, Schroeder (1999: 93) argues, is due to the fact that adjectives always appear in a classifying predicate NP that expresses individual class membership (see above), rather than e.g. status:⁶

Thus except for contexts in which being a 'young poet' or being a 'famous poet' may be an established status, the modification of a noun by means of an attribute can be expected to trigger the use of the article. This is because it expresses membership in a subset of a class—which presupposes membership in a class, thereby requiring [indefinite *bir*].

He then adds that (*ibid.*) 'we can assume that this applies for predicative as well as for non-predicative modified nouns'.

To sum up, there is evidence to suggest that in all cases indefinite *bir* serves to turn a transnumeral referent into an individual (or, rather, singleton set). If 'indefinite' *bir* is indeed a nominal aspect marker rather than an indefinite article, its position in the NP is precisely what one might expect on the basis of the *Principle of Scope*: right next to the noun.

10.2.1.2.5. Conclusion

In languages without a class of adjectives the nominal aspect marker always appears next to the noun, the only exception being Hixkaryana. This is probably due to the fact that in Hixkaryana the grammatical status of the element in question (*komo*) is not clear. It is called a particle, which in the case of Hixkaryana means that it does not seem to have a fixed position, is phonologically a bound form (like a clitic), but morpho-syntactically behaves like a free constituent (Derbyshire 1979: 83).

In languages with both a free nominal aspect marker and adjectives we also find one counter-example to the hypothesis that constituents that together make up the quality layer in the underlying representation always form a continuous sequence in the actual linguistic expression. In Guaraní, another indigenous language from South America, the nominal aspect marker may appear separated from the noun and the adjective. As in the case of Hixkaryana there are strong indications that this is only an apparent counter-example: the grammatical status of the two elements in question is far from clear. They only occur after the noun (where constituents have no fixed position), preferably after all other modifiers, and one of the elements under consideration can also modify the verb.

(p.322) It was also argued that the so-called indefinite article in Turkish *bir* (which occurs in between the adjective and the noun) is perhaps better analyzed as a nominal aspect marker, in which case its position can be explained in terms of the *Principle of Scope*. Nevertheless one could ask why in a language such as Turkish (i.e. with adjectives and a free nominal aspect marker on the same side of the head noun) the adjective is not contiguous with the noun. Although according to the *Principle of Scope* there is no ordering preference as regards the adjective and the aspect marker, there would be a preferred pattern according to the *Principle of Head Proximity*, namely with the adjective adjacent to the noun. The reason why we do not find noun-adjective adjacency here might be attributed to the observation that (as we saw before) the *Principle of Head Proximity* seems to have a greater impact on the syntax of possessor NPs and relative clauses, i.e. embedded domains, than on adjectives, which are non-embedded domains (9.3.3.5). Recall also, however, that according to Krámský indefinite *bir* only appears after emphasized adjectives (section 9.2.1.3).

10.2.2. Nominal aspect on adjectives

It was mentioned at the beginning of this section that, in most languages that have a nominal aspect operator (ω_1), this category is expressed as an *inflectional* category on the noun, but in a few languages the nominal aspect marker may also occur on the adjective. Since this is not what one might expect of an inflectional category of the noun, I will discuss such cases below.

10.2.2.1. Abkhaz

It was noted earlier (sections 2.2 and 4.2.1.2) that Abkhaz is not a clear-cut case in that only non-human nouns have certain characteristics of set nouns (see also sections 2.2.1 and 2.2.2). Hewitt (1979: 223) writes that the so-called plural suffix *-k^oa*, which is used mostly with non-humans, may also appear on the adjective. On the assumption that *-k^oa* is a collective aspect marker (rather than a plural marker) one could hypothesize that the adjective is attracted to the head noun under the influence of the *Principle of Head Proximity* (see above), but this would not explain why, according to Hewitt, the suffix may optionally appear on the noun as well (PL = collective aspect in my terms):

Abkhaz (Hewitt 1979: 223)

(42) a-la(-k ^o a) bzéya-k ^o a
ART-dog-(PL) good-PL

According to Arie Spruit (personal communication), however, adjectives form a compound with the noun and the so-called ‘plural suffix’ can only appear after the nominal compound as a whole. This would also fit the description of Abkhaz ([p.323](#)) and the other West Caucasian languages (Abaza, Adyghe, Kabardian, and Ubykh) as poly synthetic languages (Testelec 1997c: 279).

10.2.2.2. Oromo

Nominal aspect markers (‘singulative suffixes’) are also found on a few adjectives in e.g. the Boraana, Orma, and Waata dialects of Oromo, but according to Stroomer (1987: 88, 1995: 47) they are all lexicalized forms. Recall that Oromo only has a minor class of adjectives (Bender et al. 1976b: 145) and it is not clear to what extent the distribution of this particular suffix includes the small class of true adjectives. To give an example, at least some of the ‘adjectives’ listed as taking the singulative suffix *-sa* or *-ca* (such as Boraana and Orma *duree-sa* ‘rich’) ‘can be translated as nouns or as adjectives’ (Stroomer 1987: 99).⁷

10.2.2.3. Conclusion

Since nominal aspect only relates to nouns, we expect only nouns to be inflected for this category. At a first glance this expectation is frustrated by Abkhaz and Oromo, but a closer inspection reveals that in Abkhaz (if we are indeed dealing with nominal aspect marking, since the evidence is not quite conclusive) the adjective is most probably part of a polysynthetic construction and that in Oromo we are dealing with lexicalized forms.⁸

10.2.3. Numerals and demonstratives

According to the *Principle of Scope* an attributive numeral—whether a grammatical (ω_{2a}) or a lexical element (τ_{2a})—should occur closer to the head noun than a demonstrative if they occur on the same side of the head noun (section 10.1), i.e. [dem num N] or [N num dem]. As far as the evidence goes this is ([p.324](#)) confirmed in the case of Abkhaz, Alamblak, Basque,⁹ Burushaski, Chukchi, Dutch, Georgian, Guarani, Hungarian, Kayardild, Ket, Korean,¹⁰ Lango, Mandarin Chinese, Nama Hottentot, Oromo (but see below on certain dialects), Imbabura Quechua, Sarcee, Tamil, and Turkish. Unfortunately no information could be obtained on the relative order of demonstrative and numeral in Galela, Hittite, Hurrian, Ika, Nahali, Nasioi, Ngiti, Nunggubuyu (but see section 8.5.1.1), and Sumerian.

In the remaining languages the problem under consideration is not always relevant; this may be due to various reasons:

1. the demonstrative and the numeral appear on different sides of the noun (Bambara, Berbice Dutch Creole, Burmese, Hmong Njua, Korean, Nung, Samoan, Tsou, Vietnamese, Wambon);
2. demonstrative and/or the numeral modifiers are only expressed as bound forms (e.g. Gude, Nivkh; cf. also section 5.2.2.1.1 on Chukchi numerals);
3. the demonstrative and/or the numeral is not part of the (simple, integral) NP: Hixkaryana (sections 5.3.2 and 6.2.1), Koasati (sections 2.2 and 5.3.1), and probably Nunggubuyu (section 8.5.1.1).

Apart from this there is the fact that in quite a few languages a demonstrative and a numeral do not normally modify the same noun; such co-occurrence restrictions are discussed in section 10.2.4.1.

The unpredicted patterns [num dem N] and [N dem num] are attested in at least the following languages: Babungo, some Oromo dialects (cf. section 9.2.3), and West Greenlandic.¹¹ If in these languages the numeral is a *grammatical* expression of cardinality (ω_{2a}) in a simple, integral NP, this would be inexplicable in the light of the word order principles proposed here. Provided that the numeral as well as the demonstrative are fully integrated constituents of the simple NP, such an ordering can only be explained if the numeral in question is an embedded, lexical expression of cardinality ([num = τ_{2a} dem N] and [N dem num = τ_{2a}]), since according to the *Principle of Domain Integrity* the preferred position of an embedded modifier is not between constituents of the matrix NP.

(p.325) 10.2.3.1. Babungo

In Babungo the order of constituents in the NP is: Noun—Adjective—possessor pronoun—Associative NP—demonstrative—numeral—prepositional NP—relative clause. Recall that the position immediately before the head noun is reserved for pragmatic purposes (section 6.2.1.2), and that the possessor pronoun immediately follows an inalienably possessed noun, possibly forming a compound with the head noun (section 9.2.2.1). What is relevant here, of course, is that the demonstrative occurs in between the head noun and the numeral. Still, this is not a true counter-example. Since numerals are expressed in the form of an embedded NP, i.e. as a quantifying satellite τ_2 (section 5.3.2), their syntactic properties can be explained as a consequence of the fact that numerals are subject to conflicting ordering preferences. We saw earlier that according to the *Principle of Scope* (and the *Principle of Head Proximity*) the preferred pattern is

[N num $_{\tau_2}$ dem], but that the second part of the *Principle of Domain Integrity* (section 8.4.2) says that constituents of the matrix domain (here i.e. [N dem]) prefer not to be interrupted by embedded domains, such as lexical expressions of cardinality.

10.2.3.2. Oromo

According to Bender et al. (1976b: 145) adjectives, numerals, and demonstratives all follow the noun in that order (*ibid.* 144), but Stroomer (1987: 59) states that in certain dialects (Boraana, Orma, and Waata) the numeral can also appear before the adjective, or after the demonstrative (see also sections 9.2.2.2 and 9.2.3 on other problems concerning the internal syntax of Oromo NPs):¹²

Oromo (Stroomer 1987: 59; S = subject form)

(43)	haattii	ijoollee	tana	afraanii	guddoo	hiyeettii
(Boraana)						
	= haad'a-tii	ijoollee	tana	afraani ii	guddoo	hiyeettii
	mother-S	children	these	four	LIN = GEN	very poor
'The mother of these four children is very poor'						

(44) minnini	arfaani	gugurdaani	kuni	haa	jige (Waata)
= mina-nini	arfaani	gugurdaa-ni	kuni	haa	jig-e
house-S	four	big-S	these:S	FOC	collapse-M.SG.PAST
'These four big houses have collapsed'					

(45) gaala	lamaani	gugurdaani	kuni	sookoo	đaķe (Orma)
= gaala	lamaani	gugurdaa-ni	kuni	sookoo	đaķ-e
camel	two	big-S	these:S	market	go-3M.SG.PAST
'These two big camels went to the market'					

Since Stroomer does not specify the degree of markedness or the circumstances under which these constructions were produced, it is not possible to explain the (p.326) syntactic variation displayed above, but it may be interesting to note that Oromo numerals have a number of peculiar properties (see also section 4.3.2 on Oromo adjectives and note 11 on demonstratives in Chapter 9). For instance, there are certain numerals (notably between ‘one’ and ‘twelve’) that may appear with the suffix *-(aa)ni*, which Stroomer (1987:107), following Owens’s suggestion (Owens 1980), believes to be related to the plural suffix *-aani*. This could mean that numerals are a special subclass of nouns. Although numerals are not categorized by Stroomer, in Arbore, a closely related language of the same family (Cushitic), the nominal status of numerals, which clearly resemble their Oromo counterparts, is fairly undisputed (Hayward 1984:208 ff.; see also e.g. Hetzron 1967 on Awar).¹³

10.2.3.3. West Greenlandic

Apart from the order [N num dem], West Greenlandic also has [N dem num]:

West Greenlandic (Fortescue 1984: 118)

(46) qimmi-t	qaqurtu-t	marluk	taakku
dog-PL	white-PL	two	those
'those two white dogs'			

(47) qimmi-t	qaqurtu-t	taakku	marluk
dog-PL	white-PL	those	two
'those two white dogs'			

It has already been mentioned in section 5.3.2 that West Greenlandic numerals are categorized as nouns (Fortescue 1984: 204, 247), which means that according to the *Principle of Domain Integrity* the preferred position of numerals is in fact in the periphery of the NP, as in (47).¹⁴ Recall also that West Greenlandic lacks a class of true adjectives (section 4.3.3); the modifier *qaqurtut* is actually a participial form, inflectable for case and number.

10.2.3.4. Conclusion

It was predicted in section 10.2.3 that only lexical expressions of cardinality (numeral nouns and verbs) do not necessarily occur between the demonstrative (p.327) and the noun in those languages where both the demonstrative and the numeral precede or follow the noun. This appears to hold true for Babungo and West Greenlandic. In the case of one language (namely Oromo) so many details about the internal structure of the NP (see section 9.2.2.2) remain unclear that it is currently not possible to decide whether this language constitutes a true counterexample to the hypotheses formulated in section 10.2.3.

10.2.4. Overview of ordering patterns in the simple, integral NP

The purpose of this section is to examine the relative order of demonstrative, numeral, adjective, and noun in those languages in which the three modifiers are always free, fully integrated constituents of the NP. If a simple, integral NP contains a demonstrative, a non-lexical numeral, a true adjective, and the head noun, only the following eight patterns are predicted to be possible by the *Principle of Scope*, also in combination with the *Principle of Head Proximity* (first hypothesis: adjective-noun adjacency; see section 9.2) and the *Principle of Domain Integrity* (section 8.4.2):¹⁵

(48) dem num A N dem A N num num A N dem A N num dem

dem num N A dem N A num num N A dem N A num dem

There are, however, quite a few languages in which the demonstrative, numeral, and/or adjectival modifier is not (always) a free constituent of the simple, integral NP, but e.g. a bound form, an embedded (phrasal) construction, an apposition, or a sentence constituent. More specifically, Sarcee and probably Abkhaz and Chukchi have no free adjectives and adjectives are lacking altogether in Burmese, Galela, Hixkaryana, Koasati, Korean, Krongo, Lango, Mandarin Chinese, Nivkh, Nung, probably Nunggubuyu, Tsou, Vietnamese, and West Greenlandic (see section 4.3). Similarly demonstratives in Gude and (probably) Nivkh are bound forms and they are possibly appositional modifiers in Nasioi, West Greenlandic, Nunggubuyu, and Samoan (at least the postnominal variant). In Hixkaryana demonstrative pronouns are not deemed to occur as adnominal modifiers. Bound, embedded, or appositional expressions of cardinality are found in Abkhaz, Basque, Chukchi, Hurrian, and Nivkh (section 5.2.2.1.1).

To the extent that the relevant word order facts were available for the languages in the sample, Table 10.1 specifies the relative order of demonstrative, numeral, adjective, and noun in those languages in which the three modifiers are free constituents in the same (simple, integral) NP.

In some cases it has not been possible to find explicit statements about the relative order of noun modifiers; this holds for Bambara, Burushaski, Hittite, Ika, Nasioi, and Ngiti. In some other languages (especially those without a written (p.328)

Table 10.1. Constituent order in the simple NP

Language	Word order in the simple, integral NP
----------	---------------------------------------

Alamblak ¹	dem num A N
Bambara	dem N num/A
Basque	num N A dem
Berbice Dutch Creole	num A N dem
Bukiyip	dem/num A N
Burushaski	dem num N A
Dutch	dem num A N
Georgian	dem num A N
Guaraní	dem num N A
Hittite	?
Hmong Njua	num + CLF N A dem
Hungarian	dem num A N
Ika	?
Kayardild ²	dem num A N
Ket	dem num A N
Kisi	?
Nama Hottentot	dem num A N
Nasioi	?
Ngalakan	dem N A
Ngiti	?
Oromo ³	N A num dem
Pipil	dem num A N
Quechua, Imbabura	dem num A N
Tamil	dem num A N
Turkish	dem num A N
Wambon	dem N num/A

(1) Alamblak numerals commonly appear between the demonstrative and the adjective but ‘may permute to any position’ in the NP (Bruce 1984: 102). This may be due to the fact that the numeral equivalents of ‘three’ and ‘four’ exhibit certain ‘phrasal characteristics’ (Bruce 1984: 102) and that speakers of Alamblak use more than one numeral system. It is, however, not clear whether the syntactic ordering possibilities of numerals holds for members of all systems or only for certain members of some systems (Bruce 1984: 102): ‘Three numeral systems coexist in Alamblak: a borrowed tally system, the borrowed New Guinea Pidgin system, and the mixed binary/quinary/vigesimal system (primitive numbers of one, two, five, and twenty). There are some further variations based on the presently used system, i.e., money counting numerals are derived from the basic numerals.’

(2) In Kayardild (Evans 1995: 235) demonstratives, numerals, and adjectives all precede the noun in that order, but ‘counting is not a traditional activity in Kaiadilt society’ (Evans 1995: 242) and the language has very few ‘number words’

(*warirra* ‘nothing’, *warngiida* ‘one’, *kiyarrngka* ‘two’, *burldamurra* ‘three’, *mirndinda* and *kiyarrngka marlda* ‘ten’ (lit. ‘two hands’)).

(3) Recall, however, that Stroomer (1987: 59) states that in certain Oromo dialects (Boraana, Orma, and Waata) the numeral can also appear before the adjective, or at the end of the noun phrase (see also section 9.2.2.2). It is perhaps interesting to add that adjectives and demonstratives, but not numerals, are ‘gender sensitive modifiers’ (Stroomer 1987: 59).

(p.329) tradition) the grammatical description did not contain any examples of an NP containing more than one or two modifiers. Furthermore, whereas in a language such as Dutch it is quite acceptable to have a noun modified by three modifiers (as in *die drie zware kisten* [dem num A N] ‘those three heavy boxes’), in other languages such expanded structures are often avoided, marked, or simply ungrammatical. This seems to hold for at least the following languages: Bukiyp, Chukchi, Hixkaryana, Kisi, Krongo, Ngalakan, Nivkh, Pipil, and Wambon. These languages are discussed in the next section.

10.2.4.1. Co-occurrence restrictions

I mentioned in the previous section that the sample contains several languages in which NPs with two or three modifiers are avoided (see also section 1.5.4.3). In their grammatical overview of Nivkh, which takes into account data from all major sources, Mattissen and Drossard (1998: 51) write that they have ‘no examples which would show whether determiners, possessors, and quantifiers can be combined, and there are only a few examples of a complex containing both a determiner and an attribute in our data’ (recall that NPs in Nivkh are poly synthetic structures):

Nivkh (Mattissen and Drossard 1989: 51; Jakobson 1971a: 80)

(49) hun-tleulan-*ti*ř

that-white-hill

Since in Bukiyp [dem/num A N] demonstratives and numerals are assigned to the same syntactic slot in the NP-structure provided by Conrad (1991: 57), I have assumed that these modifier categories do not normally occur simultaneously in the same NP.¹⁶

Derbyshire (1979: 44) writes that in Hixkaryana (note that I have omitted the paragraph about possessive modifiers):

[p]reposed modifiers are infrequent in noun phrases, being restricted to numerals and the nominal form anaro ‘another’, both of which have other, preferred, functions. Numerals more often occur as sentential adverbs, related to the noun phrase, but separate from it. The nominal, anaro, usually occurs in a separate noun or postpositional phrase, in para- tactic relationship with the noun it modifies

Postposed particles may be any of the three sub-classes of postpositional particles ...: heno ‘now dead’, kom(o) ‘COLLECTIVE’ and ymo ‘AUGMENTATIVE’ are modifying particles; haxa ‘CONTRAST’ and ryhe ‘EMPHATIC’ are discourse particles; and hati ‘HEARSAY’ is a verification particle sequence

There is one other type of construction—the equative sentence—that resembles a noun phrase. It is the construction that handles some types of what in other languages are adjective-noun and demonstrative-noun relations. The possibilities of expansion of these constructions and their range of syntactic functions are, however, of the sentence kind rather than the phrase.

(p.330) The pattern given for Ngalakan in Table 10.1 lacks a numeral, because Merlan (1983: 71, 83) treats the two Ngalakan numerals (*wangiñ* ? ‘one’ and *yapan* ? ‘two’) as adjectives, which normally follow the noun (ibid. 80). She also writes (ibid.; see also section 1.5.4.1):

In general, the Ngalakan NP exhibits a fairly loose sort of structure. It is possible for constituents of what could be considered the ‘same’ NP to be separated from each other by other clausal constituents, or for many NP constituents having the same referent to be strung together in a fairly loose sort of appositional structure.

In Chukchi, Kisi, Krongo, and Pipil there also seems to be a general avoidance of having multiple modifiers in one NP; at least the text materials of these languages rarely contain cases of a noun modified by e.g. a demonstrative and a numeral or a demonstrative and an adjective, let alone a noun simultaneously modified by a member of all three modifier categories. For example, Tucker Childs (personal communication) states that his Kisi informant ‘did not like numbers and demonstratives together’.

It is nevertheless often possible to get the construction one is interested in (and that is not attested in the transcripts of spoken language) by elicitation, as was the case with this example from Chukchi:

Chukchi (Vladimir Nedjalkov, personal communication)

(50)	əŋgena-t	ngeroq	n-ilg-ə-qine-t	qora-t
	this-PL	three	ADJ-white-ə-3-PL	deer-PL
‘these three white deer’				

This construction is however only possible when the NP is in the absolute case (intransitive subject, direct object); in all other cases numeral and adjective are incorporated (and the demonstrative was simply not included when informants translated the NP):

Chukchi (Vladimir Nedjalkov, personal communication)

(51)	ətlon	ga-twetcha-twa-len	ga-ngeron-elg-ə-qaa-ma
	he	PERF-stand_up-be-3SG	COM-three-white-ə-deer-COM
‘He stood next to these three white deer’			

According to Lyle Campbell (personal communication) there would be nothing ungrammatical about an NP with all three modifiers in Pipil (in which case the order would be [dem num A N]), but at the same time his intuition tells him this is just never used.

It was pointed out in Chapter 1 (section 1.5.4.3) that speakers of certain languages prefer to distribute many modifiers over several NPs rather than have them all in one NP. This is also true for Wambon. Although it is theoretically possible to produce a construction that has all three modifiers in it, such NPs are not attested in actual spoken discourse. Instead speakers of Wambon would (p.331) distribute constituents over two juxtaposed NPs, as in: Wambon (Lourens de Vries, personal communication)

(52)	ev-o	kap	ambalopkup	ev-o	kap	kaimombalin
	that-CN	man	five	that-CN	man	good
'those five good men'						

10.2.4.2. Orderings patterns in the noun phrase

The data in Table 10.1 show that only some of the possible orders are actually attested in the sample:

(53)	dem num	Alamblak, Dutch, Georgian, Hungarian, Kayardild, Ket, Nama Hottentot, Imbabura
	A N	Quechua, Pipil, Tamil, Turkish
	dem num	Burushaski, Guaraní
	N A	
	dem A N	
	num	
	dem N A	Bambara?
	num	
	num A N	Berbice Dutch Creole
	dem	
	num N A	Basque, Hmong Njua
	dem	
	A N num	
	dem	
	N A num	Oromo
	dem	

The fact that some of these patterns are not attested is at least partly due to the relatively small size of the sample. We know, for instance, that the pattern [dem num N A] is also found in the Romance languages. And according to Hawkins (1983: 119) the pattern [dem N A num] is found in Kabardian and Warao and the pattern [N A num dem] in Selepet and Yoruba, although it remains to be seen if we are really dealing with simple, integral NPs here. If we include these other languages, the result would be as follows (languages in square brackets are from Hawkins):¹⁷

(54)	dem	Alamblak, Dutch, Georgian, Hungarian, Kayardild, Ket, Nama Hottentot, Imbabura
	num A	Quechua, Pipil, Tamil, Turkish [English, Finnish, Hindi, Maung]
	N	
	dem	Burushaski, Guaraní [French, Italian, Spanish]
	num N	

A	
dem A	
N num	
dem N	Bambara? [Kabardian, Warao]
A num	
num A	Berbice Dutch Creole
N dem	
num N	Basque, Hmong Njua, [Easter Island, Indonesian, Jacaltec, Maori, Welsh]
A dem	
A N	
num	
dem	
N A	Oromo [Selepet, Yoruba]
num	
dem	

(p.332) This still leaves a gap for at least two patterns: [dem A N num] and [A N num dem].¹⁸ These non-attested patterns seem to indicate that (in the simple, integral NP) a numeral only appears after the noun if the adjective does so too.¹⁹ Thus, there seem to be implicational universals concerning the relative order of noun modifiers in the simple NP (if N num, then N A) for which there does not seem to be a good explanation —unless of course numerals are treated as a subclass of adjectives.

10.3. The position of embedded modifiers

The *Principle of Scope* says that an embedded domain (like any other expression of an operator or satellite) must occur next to the element(s) it has in its scope. As shown earlier (section 10.1), the following patterns are all permitted by the *Principle of Scope*:²⁰

(55)

- a. $\omega_{2b} \omega_{2a} \omega_1 N \tau_1 \tau_{2a} \tau_{2b}$
- b. $\tau_{2b} \tau_{2a} \tau_1 N \omega_1 \omega_{2a} \omega_{2b}$
- c. $\omega_{2b} \tau_{2a} \omega_1 N \tau_1 \omega_{2a} \tau_{2b}$
- d. $\tau_{2b} \omega_{2a} \tau_1 N \omega_1 \tau_{2a} \omega_{2b}$
- e. $\tau_{2b} \omega_{2b} \tau_{2a} \omega_{2a} \tau_1 \omega_1 N$
- f. $N \omega_1 \tau_1 \omega_{2a} \tau_{2a} \omega_{2b} \tau_{2b}$
- g. $\omega_{2b} \tau_{2b} \omega_{2a} \tau_{2a} \omega_1 \tau_1 N$
- h. $N \tau_1 \omega_1 \tau_{2a} \tau_{2a} \tau_{2b} \omega_{2b}$

Since I have found no instances in which the same noun is modified by both an embedded (lexical) and a non-embedded (grammatical) expression of cardinality (ω_{2a} and τ_{2a}), I will confine myself here to combinations of grammatical and (p.333) lexical expressions of the notion ‘location’ in the NP (localizing operators and satellites ω_{2b} and τ_{2b} , respectively). In case a localizing operator and satellite appear on the same side of the noun, there is no preferred order in terms of the *Principle of Scope*, so we may expect all four possibilities to occur (as to languages in which the position of localizing satellites G and Rel in the NP violates the *Principle of Scope*, as well as the *Principle of Domain Integrity*, I refer to section 9.5 of the previous chapter on Head Proximity):

(56)

- a. $\tau_{2b} \omega_{2b} \dots N$
- b. $N \dots \omega_{2b} \tau_{2b}$

- c. $\omega_{2b} \tau_{2b} \dots N$
- d. $N \dots \tau_{2b} \omega_{2b}$

The orders $[\tau_{2b} \omega_{2b} \dots N]$ [$\omega_{2b} \tau_{2b} \dots N$] are attested in quite a few languages (see section 9.5); both of them occur in Turkish:

Turkish (Schroeder 1999: 32)

-
- (57) a. oyna-yan bu küçük çocuk (or bu oyna-yan küçük çocuk)
 play-S.Part this small child
 'this small child who is playing'
-

The other two orders are also found. The pattern $[N \dots \omega_{2b} \tau_{2b}]$ is attested in, for example, the Celtic languages, where the order of postnominal modifiers is [... N A dem G Rel] (Tallerman 1997: 37).

The pattern $[N \dots \tau_{2b} \omega_{2b}]$ is attested in e.g. Nung, where the demonstrative (ω_{2b}) follows the relative clause (τ_{2b}).²¹

Nung (Saul and Freiberger Wilson 1980: 16)

-
- (58) vahng khị tú bê tê
 boy ride CLF goat that
 'that boy (who was) riding the goat'
-

(p.334) In languages in which the localizing satellite τ_{2b} is between the noun and the localizing operator ω_{2b} (as in the patterns $[\omega_{2b} \tau_{2b} \dots N]$ and $[N \dots \tau_{2b} \omega_{2b}]$) there is the danger of potential ambiguity in that the operator can be interpreted as a constituent of the localizing satellite (observe that such a pattern would be preferred in terms of *Head Proximity* but not in terms of *Domain Integrity*, see section 9.5). This is explicitly stated in the case of Nung (Saul and Freiberger Wilson (1980: 37): 'When there is an embedded noun phrase or clause within the main noun phrase, it is often ambiguous as to whether the Demonstrative is modifying the embedded phrase or the main phrase.' Haspelmath (1993: 269) suggests that this kind of ambiguity is deliberately avoided in Lezgian (see also note 38 on Udmurt in Chapter 9): 'The Genitive NP generally precedes all other modifiers because otherwise those modifiers could be taken as modifying the Genitive noun rather than its head.' For this reason languages often allow the demonstrative to precede or follow the localizing satellite. In the sample this variation is attested in e.g. Abkhaz, as well as in Burushaski (Lorimer 1935–8: i. 404), Lango (Noonan 1992: 156), Tamil, and Turkish (see above). Hewitt (1979: 61) writes that in Abkhaz '[demonstratives may precede relatives if there is no possibility of the demonstrative being interpreted as a constituent of the relativized clause'].

Abkhaz (Hewitt 1979: 61)

-
- (59) wəy a-xàc 'a də-z-bà-z à-jyab
 that ART-man him-who-see-NonF ART-girl
 'the girl who saw that man'
-

-
- (60) a-xàc 'a də-z-bà-z wəy à-jyab
 ART-man him-who-see-NonF that ART-girl
-

'that girl who saw the man'

In Tamil possible ambiguity can also be avoided by means of a pause and an intonation break between the demonstrative and the relative clause (Asher 1982: 66).

10.4. Conclusion

Although the word order principles discussed in the last three chapters go a long way in accounting for attested patterns, there is still much that needs to be explained. In the first place we need to find an explanation why patterns that are expected or predicted to be possible (such as the orderings [dem A N num] and [A N num dem]) do not seem to be used in any natural language (assuming, as I mentioned earlier, that adjectives and numerals constitute different word classes).

Secondly, it has become clear that we need at least one more principle to account for the attested word order patterns in the NP besides the *Principle of Domain Integrity*, the *Principle of Head Proximity*, and the *Principle of Scope*. (p.335) This principle must account for the fact that in some languages emphasized constituents may occur in a *special position* (e.g. the NP-initial position in Babungo, Hungarian, Tamil; the postadjectival/postnominal position in Modern Greek). To be sure, such a principle has already been formulated by Dik as the *Principle of Pragmatic Highlighting* (Dik 1997: 403), who also noted (*ibid.* 430) that this principle is much less influential in the domain of the NP than in that of the clause: 'Constituents with special pragmatic functionality ... are preferably placed in 'special positions', including, at least, the clause-initial position.'

Thirdly, if we also want to take into account ordering phenomena in complex NPs (involving embedded modifiers), we need to take into consideration the role that heaviness or complexity plays in syntax (cf. Hawkins 1983: 98 f.; also Hawkins 1994).

Fourthly, *ambiguity* may be the cause for certain deviations in ordering patterns as preferred in the light of general ordering principles and it is also far from clear why the same principle can have a different impact on word order patterns in different languages, even without considering the role of other principles. Last but not least, it is evident that we not only need more explanations, but also more reliable data from a wide variety of the world's languages.

Notes:

(1) Nor does the *Principle of Scope* say anything about the ordering of stacked adjectives; see note 7 in Chapter 7.

(2) See also section 9.5 in the chapter on Head Proximity on the syntax of embedded domains.

(3) Recall that Sarcee is a doubtful case, since in this language 'adjectives' are either treated as part of a compound (comparable with English *blackbird*) or as verbal elements heading a relative clause (Cook 1984: 67).

(4) The particle *heno* means 'quantity, set of if used in connection with non-human referents (Derbyshire 1979: 83).

(5) See Schroeder (1999: 89 f.) for an analysis of those cases in which indefinite *bir* is not always used to introduce a referent for further reference. One of these cases involves modified nouns, which are discussed below.

(6) In this context (i.e. of non-verbal predication) the notion of 'status assignment' is taken from Hengeveld (1992a: 76), citing Dik (1980a: 98), and refers to bare nominal predicates designating 'membership of some established functional, professional or ideological group'.

(7) To give another example, Bender et al. (1976b: 134) write that the literal translation of *nam-icci sore-sa* [man-SING rich-SING] reads: '(the) man (the) rich one' = 'the man is rich'. In this example (with a predicative adjective, if it is indeed an adjective) the singulative aspect marker seems to serve as a derivational rather than an inflectional suffix. See also Chapter 4, note 26 on the relationship between markers for agentivity and singularity.

(8) In the Asmat dialect described by Voorhoeve (1965: 138) the suffixes *-nakáp* and *-nakás*, which mainly express the diminutive as well as individual and collective aspect, respectively, may also appear on adjectives. But on these elements they have a different meaning, namely ‘much, to a high degree’. Compare *jiníc* ‘slack’ vs. *jinícnakíp* ‘very slack, weak’ or *nemcén* ‘angry, pugnacious’ vs. *nemcénnakás* ‘furious, very pugnacious’. Voorhoeve (*ibid.*) writes that is not clear if the two suffixes are always interchangeable.

To give another interesting example in this context: in Gambian Mandinka the so-called plural suffix *-lu* also appears on the adjective, if there is one, as in *téeri bétoo-lu* [friend good-LU] ‘good friends’. Here one might suspect that the unclear morpho-syntactic status of the ‘plural’ marker plays a role, for *-lu* has certain characteristics ‘which could justify writing it as a separate word’ (Rowlands 1959: 38). The properties in question are that it has a separate accent of its own and that it sometimes follows the ‘sentence modifier’ *le* (probably a focus marker), which it usually precedes. If, indeed, it is a free element, then the order N A(-)*lu* would be in accordance with both the *Principle of Head Proximity* and the *Principle of Scope*.

(9) The numerals ‘one’ and in some dialects ‘two’ follow the noun, other numerals are in the Prefield (Chapter 5).

(10) See Chapter 4 on the formal expression of numerals in Korean.

(11) The examples in the grammar of Krongo (Reh 1985: 249–50) indicate that the order is [N num dem] but Reh’s unpublished field notes (personal communication) contain an NP with the sequence [N dem num A] (cf. also Reh 1994: 236 on Deiga, a language that is closely related to Krongo). Recall, however, that numerals (as well as adjectives) are actually verbs in Krongo (Reh 1985: 242 f.), i.e. embedded domains in the NP, their positional properties can be explained as a consequence of the fact that lexical expressions of cardinality are subject to conflicting ordering preferences. According to the *Principle of Scope* and the *Principle of Head Proximity* the preferred pattern is [N num dem], but according to the *Principle of Domain Integrity* (section 8.4.2) the preferred pattern is [N dem num].

(12) On the Harar dialect of Oromo described by Owens the usual position for the numeral is also between the adjective and the demonstrative (N A num dem; Owens 1985: 86).

(13) Compare also Greenberg (1978b: 279) in connection with the expression of cardinality in Semitic languages: ‘In SIDAMO, an EASTERN CUSHITIC language, 50–90 are the plurals of 5–9. In the light of these examples, the SEMITIC expression of 30, 40, etc. as the plurals of 3, 4, respectively which has puzzled Semitists, becomes clear. Originally, 20 was the dual of 10, but has everywhere except in ETHIOPEAN SEMITIC, in analogy to the other decades, become the plural of 10, whereas in ETHIOPIC the dual was generalized from 20 to the remaining decades.’

(14) Fortescue (1984: 118) also writes that in West Greenlandic scope is ‘the principal determinant of successive ordering’ in the NP and that ‘all [other] modifiers between a head and demonstrative are inside the scope of the latter’s definiteness, whereas further modifiers to the right modify the demonstrative itself (subsuming the material to its left)’. In other words, when the numeral follows the demonstrative the numeral is deemed to modify the demonstrative rather than the noun. Since the demonstrative can also occur by itself, this could indicate that the demonstrative together with the numeral are not part of the integral NP containing the head noun but instead constitute an appositional modifier construction (see also Chapter 6 on West Greenlandic demonstratives).

(15) As to the second hypothesis based on the *Principle of Head Proximity*, concerning the pre- or postnominal appearance of demonstratives, numerals, and adjectives, see section 9.3.

(16) Recall that an adjective may also follow the noun in Bukiyp (section 4.3.4).

(17) The patterns in question do occur in non-simple and/or non-integral NPs (which fall outside the scope of this book). For example, Korean also has the pattern [dem A N num], but in this language the numeral is in fact an apposed numeral classifier phrase (see Chapters 1 and 5).

(18) Gude could have been an example of a language with the order [A N num dem], were it not for the fact that the demonstrative is a suffix, which may also appear on the noun (Hoskison 1983: 45). Hawkins (1994: 318) suggests that adjectives are not preferred in NP-initial position, because they cannot ‘construct NP’ (in terms of a Chomskyan framework), as opposed to a demonstrative or a head noun. But if a language indeed has a distinct class of adjectives, I fail to see why the occurrence of an adjective in NP-initial position cannot serve as a good indication for the parser that it is now dealing with a noun phrase.

(19) A possible candidate is e.g. Zande which has [A N num]. The demonstrative consists of two elements, one of which occurs in the Prefield whereas the second part appears in the Postfield, e.g. *gu kúmbú rè* ‘that man there’ vs. *gi kúmbá rè* ‘that man here’; Tucker 1959: 143; also Gore 1926: 37). Further details are lacking unfortunately.

(20) We saw in Chapter 6 that there is no one-to-one relationship between form and function in that the same construction may be used to express different modifier functions. For instance, a relative clause may function as a qualifying satellite τ_1 (as in Galela), a quantifying satellite τ_{2a} (as in Samoan), or a localizing satellite τ_{2b} . This means that, dependent on its position in the underlying structure, a relative clause will have different ordering preferences according to the *Principle of Scope*. I have few straightforward examples to illustrate this, but Mandarin Chinese may be a case in point (also e.g. Boumaa Fijian, in which relative clauses expressing cardinality precede the head noun, whereas non-quantifying relative clauses normally follow). Descriptive relatives, i.e. relative clauses that serve as qualifying satellites in the NP (τ_1) in that they specify a non-identifying property, occur in between the demonstrative and numeral (plus classifier), which is in accordance with the *Principle of Scope* (and the *Principle of Head Proximity*, but not with the *Principle of Domain Integrity*).

Mandarin Chinese (Hashimoto 1971: 24–5)

(1) nei	ge	dai	yanjing	de	xiaohair	hen	pang
that CLF wear glasses NomP child very fat							
'the child, who wears glasses, is very fat'							

Restrictive relative clauses, i.e. relative clauses that serve as localizing satellites in the NP (r_{2b}), occur before the constituents of the matrix NP, thus adhering to the *Principle of Domain Integrity* as well as the *Principle of Scope* (cf. section 9.5).

Mandarin Chinese (Hashimoto 1971: 24–5)

(2) dai	yanjing	de	nei	ge	xiaohair	hen	pang
wear glasses NomP that CLF child very fat							
'the child who wears glasses is very fat'							

(21) Nung is one of the languages in which the numeral classifier phrase precedes the noun, which usually implies that the classifier has certain other peculiar functions (see section 5.2.2.1.3). This is also true in the case of Nung: ‘General classifiers ... may substitute for Noun Head in a noun phrase, but are usually accompanied by a numeral, modifier, a possessive, or a demonstrative’ (Saul and Freiberger Wilson 1980: 25). The classifier *tú* is used with children’s names, and nouns which designate animate beings, which include animals, humans, and spirits (ibid. 26).

11 Epilogue

11.1. Introduction

This concluding chapter is both a summary of and a comment on the contents of this book in that it provides an overview of the main points while at the same time serving as a platform for some qualifying remarks regarding certain issues addressed in previous chapters. This is also a good place to emphasize, once again, that

- this book is mainly restricted to lexical NPs which are headed by an underived noun and which are used to refer to a single, discrete, first order (i.e. spatial) entity;
- the immediate referent of an NP is not an entity in the real world, but rather a mental construct, whose properties need not coincide with the properties of its ontological correlate in the external world (if it exists).

11.2. Aims

I stated at the outset that it is one of the aims of this book to give a cross-linguistic overview of the constituents that together make up the integral noun phrase. Although most NP constituents were discussed in some detail (articles, demonstratives, numerals, adjectives, nouns, adnominal possessor constructions, relative clauses), there are also some modifier categories that received relatively little or no attention in this book. The main reason for this omission is simply that there are still quite a few gaps in our knowledge of noun modifiers from a typological point of view. This holds in particular for the various elements that, in one way or another, are concerned with systems of nominal classification (Craig 1986b; Harvey and Reid 1997; Senft 2000b); hence Chapter 3 is largely encyclopedic in nature. Nevertheless, it is perhaps interesting to note that Craig (1992; see also Grinevald 2000) has argued that the various kinds of classifiers can be associated with different layers in the model of the noun phrase discussed in Chapter 7.

Another phenomenon that deserves more attention is the distinction I made between two major kinds of modifier categories: descriptive modifiers and discourse modifiers (sections 7.6 and 7.7). Descriptive noun modifiers specify properties of the referent of the NP in terms of some of Aristotle's basic categories: the *ποιόν* 'how it is' (Quality), the *ποδόν* 'how much/many it is' (Quantity), and the *ποῦ* 'where it is' (Location). Discourse modifiers, on the other hand, are (p.337) concerned with the referential status of entities (objects, events) in the world of discourse. I will return to this distinction below.

The second aim of this book is to provide a typological adequate model of the NP in the theoretical framework of Simon Dik's *Functional Grammar* (Dik 1997). In Chapter 7 I have proposed an underlying structure that consists of four hierarchically ordered 'layers':

- the quality layer, which contains the head noun and which accommodates modifier categories that only relate to the property that is designated by the noun: nominal aspect markers and (typically) adjectives;

- the quantity layer, which contains the quality layer and which accommodates modifier categories having to do with number (singular, plural) and cardinality (one, two, etc.). Indefinite quantifiers (such as ‘some’ and ‘many’) and ordinal numerals also belong to the group of noun modifiers that were more or less ignored in this study (see above). I should point out, however, that I regard the former but not the latter as quantifying modifiers. Although ordinal numerals were briefly mentioned in Chapter 5 on quantifying modifiers, I would rather treat them as localizing modifiers (Chapter 6), since their main function is to specify the location (position) of an entity in a linearly ordered series of similar entities. Interestingly there are many languages in which ordinality is expressed in terms of spatial notions (section 5.2.2.2.);
- the location layer, which contains the quantity layer and which accommodates modifier categories specifying properties concerning the location of the referent, such as demonstratives, possessive modifiers, and relative clauses;
- the ‘referential’ or ‘discourse’ layer, which contains the location layer and which accommodates modifiers that provide the addressee with information about the referent of the NP as a discourse entity. For example, the use of a definite article typically indicates that the referent of the NP has been introduced before in the world of discourse and is therefore an identifiable entity (but note that the definite article may also be used when the occurrence of an entity in the world of discourse is presupposed; Chapter 7).

Furthermore, it was argued that NPs and clauses can be analyzed in a similar fashion in that (at some level of abstraction) they have the same kind of underlying structure, accommodating the same kind of modifier categories (more on this below).

Lastly this book gives an account of the internal syntax of NPs in terms of only a few universal ordering principles, and this is the topic of Chapters 8–10. It is argued in these chapters that the order of constituents in the NP mainly results from the interaction of three general word order principles, which are in fact all elaborations of a principle formulated by Behaghel (1932: 4), according to (p.338) which elements that belong together semantically tend to occur together syntactically:

- the Principle of Domain Integrity;
- the Principle of Head Proximity;
- the Principle of Scope.

It was established that, in addition, we need Dik’s principle of *Pragmatic Highlighting* to account for the fact that many languages permit an emphasized constituent to appear in a special position, possibly one that would not be permitted by (the interaction of) the three principles mentioned above. I will return to the general ordering principles below.

11.3. Seinsarten

I argued in Chapter 2 that semantically nouns can be analyzed in much the same way as verbs. In verb semantics, studies concerning lexical properties that are relevant for the kind of temporal entity the verb helps to define (e.g. a state, an action, an achievement) are generally subsumed under the general heading of *Aktionsart* (‘mode of action’). Similarly, the investigation of semantic properties of nouns that are relevant for the kind of spatial entity they help to define (e.g. a mass, a collective, a singular object) can

be referred to as the study of *Seinsart* ('mode of being'). I proposed a classification of six basic *Seinsarten* in terms of two spatial features (Shape and Homogeneity) (Fig. 11.1).

Nouns characterized by the feature –Shape have in common that they are not marked for number distinctions and that they require a classifier when they are modified by a numeral: sort nouns take a sortal (or numeral) classifier, mass nouns occur with a mensural classifier, and general nouns appear with a type of classifier (coined *general classifiers* in Chapter 2) that seems to be a combination of sortal and mensural classifiers. Set nouns (which are also transnumeral),

Space	–Homogeneity	+Homogeneity
–Shape	general noun	
	sort noun	mass noun
+Shape	set noun	
	singular object noun	collective noun

Fig. 11.1. Cross-linguistic classification of major nominal subcategories (or *Seinsarten*)
(p.339) singular object nouns, and collective nouns, on the other hand, can be in a direct combination with a cardinal numeral.

It turned out that cross-linguistically four noun types are used to refer to a discrete spatial object (such as a knife or a dog): general nouns, sort nouns, set nouns, and singular object nouns. It will be recalled that the direct referent of an NP is a mental construct (rather than an entity in the real world), and that this is the reason why it is possible to have certain (aspectual) discrepancies between properties of referents and properties of the ontological counterparts of these referents in the real world (if they exist). Whereas singular object nouns can be said to 'agree' with the ontological facts, the other three nominal subcategories are all in a sense semantically 'underspecified' from a Eurocentric perspective, which of course at the same time makes them more flexible for referential purposes. For example, a singular object noun can only be used in its unmarked form to refer to a singular object. So if reference is made to a non-singular entity, this *must* be explicitly indicated in the case of a singular object noun (singular 'book' vs. plural 'book-s'), but not in the case of a set noun, a sort noun, or a general noun, which are all transnumeral (section 2.3). Since predicate nouns and incorporated nouns do not denote spatial properties (i.e. they have no *Seinsari*), they cannot be specified in terms of the features Shape and Homogeneity (section 2.5).

11.4. Nominal classification

In this book I have used a rather broad definition of noun classification. I consider a language L to have a system of nominal classification when the grammar of language L contains one or more (semantic, syntactic, morphological, or phonological) rules that explicitly refer to a significant subset of all the nouns in language L.¹ Since this

definition covers both overt and covert systems of nominal classification I have considered systems that involve some kind of overtly expressed classifying element as well as systems in which nominal subclassification is manifested in some non-morphological fashion.

Thus Chapter 3 is concerned with a wide variety of ways in which nominal classification is manifested in the languages of the world. I distinguished between reflections of noun classification inside and outside the NP and where possible I tried to separate systems of noun classification proper (gender systems) from instances of apparent noun classification, in which properties of the referent rather than of the noun are at issue (classifier systems). There is an important (p.340) difference between the two systems in that classifiers (as opposed to genders) ascribe ‘discoursally relevant properties to the otherwise semantically vague noun’ (Foley 1997: 245).

11.5. Qualifying modifiers

In Chapter 2 the notion of *Seinsart* was proposed as the nominal/spatial counterpart of verbal/temporal *Aktionsart*. Parallels between the nominal and the verbal domain were further explored in Chapter 4, which is concerned with qualifying modifiers. I argued that nominal meanings can be further specified as to the way they are represented in the spatial dimension by nominal aspect markers (symbolized by ω_1 in the underlying structure of the NP), just as verbal meanings can be further specified as to the way they are represented in time by verbal aspect markers. Nominal aspect markers can be found on set nouns, where they modify the *Seinsart* of the noun, indicating what *kind* of set is being referred to by the speaker: a singleton set (i.e. a set that consists of a single individual) or a collective set. In the literature the elements that I consider to be nominal aspect markers (i.e. singulative and collective aspect markers) are commonly characterized as number markers, despite the fact that there are important differences between them. Apart from the difference in meaning, nominal aspect markers are usually optional and only appear on set nouns, i.e. transnumeral nouns that can be directly modified by a numeral. By contrast, number markers are obligatory and only appear on singular object and collective nouns. Additionally it was found that regular cases of ‘number’ discord between the predicate and argument are only attested in languages with set nouns (section 4.2.1.2).

In a previous publication (Rijkhoff 1990b; see also Grinevald 2000: 77) I suggested that sortal classifiers are also instances of nominal aspect markers, because they are deemed to ‘individualize’ the property denoted by a sort noun (Lyons 1977: 462; Grinevald 2000: 74). Although the sample contains an example of a language in which sortal classifiers serve as true nominal aspect markers (Hmong Njua; section 4.2.1.2), there are two reasons why I am now more reluctant to make such a general statement. One reason is that postnominal classifier phrases are very often appositional constructions, in which case the sortal classifier would not be an integral modifier of the NP containing the head noun. The other reason is that prenominal sortal classifiers are usually in a more advanced stage of grammaticalization and have acquired other, more discourse-like functions (such as marking definiteness or topicality; section 5.2.2.1.3), which makes it difficult to treat them as ‘pure’ nominal aspect markers.²

(p.341) As in the case of quantifying and localizing modifiers discussed in Chapters 5 and 6, a distinction was made between grammatical and lexical modes of expression and the second part of Chapter 4 was devoted to *lexical* modifiers expressing

qualitative notions on the NP (symbolized by τ_1 in the underlying structure of the NP). Since adjectives are typically used for this purpose, I concentrated on this word class. It was emphasized, however, that not all languages have a class of adjectives and that the same qualifying function can also be fulfilled by modifier constructions involving other word classes (such as abstract nouns heading an adnominal NP or a stative verb heading a relative clause).

One of the major issues in the section on adjectives concerns the question whether it is possible to explain why some languages employ a distinct class of adjectives whereas others do not. Although a full explanation is still lacking, I have argued that a necessary condition for a language to have adjectives has to do with *Seinsart*. A distinct class of adjectives is only attested in languages that use set nouns or singular object nouns (i.e. nouns with the feature + Shape) to refer to a single, discrete, spatial entity.

11.6. Quantifying modifiers

Chapter 5 is concerned with plural number marking and grammatical and lexical expressions of cardinality. Only individual object nouns and collective nouns were considered to have plural number marking in the proper sense. In all other instances so-called number markers were analyzed as aspect markers (in the case of set nouns; section 4.2.1.2) or indefinite quantifiers (in the case of sort nouns; cf. section 5.2.1.3). Recall also that in some cases the category of number (if it is that) is marked on the NP rather than on the noun (section 2.2).

In the section on cardinality a distinction was made between grammatical and lexical modes of expression (symbolized by ω_{2a} and τ_{2a} respectively in the underlying structure of the NP). Numeral verbs and nouns were regarded as lexical expressions of cardinality (*quantifying satellites*). They take the form of embedded structures in their modifying function, thus turning the NP into a complex expression. In all other cases numerals were treated as grammatical expressions of cardinality (*quantifying operators*) in the NP.

(p.342) It was also established that the function of a numeral depends on the *Seinsart* of the head noun. It seems to serve as a multiplier in the case of individual object and collective nouns, whereas in the case of a set noun it specifies the size of the set (section 5.2.1). In languages with sort nouns, the cardinal numeral is typically in a direct combination with the classifier rather than the head noun, especially when the classifier phrase (numeral + classifier) has an appositional relationship with the phrase containing the head noun (section 5.2.2.1.3).

11.7. Localizing modifiers

Localizing modifiers relate to the location of the referent in the world of discourse. The *localizing operators* that are discussed in Chapter 6 are articles and demonstrative pronouns (symbolized by ω_{2b} in the underlying structure of the NP). It is already announced in the same chapter (note 2), however, that the function of articles is also discussed in Chapter 7, where it is argued that articles are better interpreted as discourse operators (symbolized by Ω_3). Rather than specifying the *location* of the referent in the world of discourse, the basic function of the (in)definite article is to indicate the *discourse status* of the referent in the world of discourse, in that it signals

to the hearer whether or not the speaker considers a referent to be identifiable in that world of discourse.

Localizing satellites (symbolized by τ_{2b}) are all those adnominal modifiers that involve lexical means to specify the location of the referent in the world of discourse. This includes relative clauses, possessor NPs, and of course phrases like ‘on the table’ in, for example, ‘Could you give me **the book on the table**, please?’

It was also established in the introduction to Chapter 6 that localizing modifiers (or, rather, their referents) are often used to license the introduction of new entities as identifiable referents by providing them with a referential ‘anchor’, the so-called *Identifying Entity*. Thus, even though the addressee may have been previously unaware of the existence of a restaurant in the Van Gogh Museum, the speaker can still refer to this restaurant as an identifiable (definite) entity in ‘Yesterday Susan and I had lunch in *the restaurant* at the Van Gogh Museum’. By using the localizing satellite ‘at the Van Gogh Museum’ the speaker provides the addressee with a referential anchor, the *Identifying Entity*, that makes it possible to infer the location of the new referent, the restaurant (whose existence is presupposed), in the world of discourse (section 6.1).

11.8. The underlying structure of noun phrases

Chapter 7 contains a proposal for the representation of the underlying structure of integral NPs. It is a semantic model that consists of four hierarchically organized (‘nested’) layers, each layer accommodating grammatical and lexical expressions (p.343) of a distinct descriptive or referential modifier category (see section 11.2 above). Since there is no one-to-one relationship between form and semantic function, the same kind of construction may be used to express different kinds of modifier categories. This holds in particular for adnominal NPs and relative clauses, both of which can occur as qualifying, quantifying, and localizing satellites (see e.g. sections 4.1 and 7.4).

Due to the restrictions mentioned at the beginning of this chapter, the current study does not consider NPs headed by nouns denoting abstract (‘love’), temporal (‘meeting’), or even higher order entities (such as ‘opinion’ or ‘question’), but it is obvious that such nouns do not always occur with the same modifiers as nouns that are used to refer to discrete spatial entities (Rijkhoff 2001). For example, events (but not spatial objects) can be modified by adjectives expressing duration (cf. ‘a brief meeting’ vs. *‘a brief knife’), although some adjectives can be used for different kinds of entities (‘a short meeting’, ‘a short knife’).

I argued that NPs and clauses (predications) can be analyzed in a similar fashion, as reflected in the simplified version of the formal representation of their underlying structures below (where Ω and ω symbolize grammatical categories and τ and T lexical modifiers in the NP, and where Π and π symbolize grammatical categories and σ and Σ lexical modifiers in the clause; the reader is referred to Chapter 7 for more details).

I mentioned that parallels in the underlying structures of NPs and clauses may be due to the fact that temporal (more complex) entities are conceived of as spatial (less complex) entities, since the latter are probably easier to deal with in human cognition. Although it is quite possible that the observed parallels between NPs and predications

are due to the metaphorical extension from the spatial to the temporal domain, there are also philosophical and scientific arguments to explain

$(\Omega_3 [\omega_{2b} [\omega_{2a} [\omega_1 N \tau_1] \tau_{2a}] \tau_{2b}] T_3)$

--

nucleus

quality layer L_1

quantity layer L_{2a}

location layer L_{2b}

discourse layer L_3

Fig. 11.2. The hierarchical structure of the NP
(p.344)

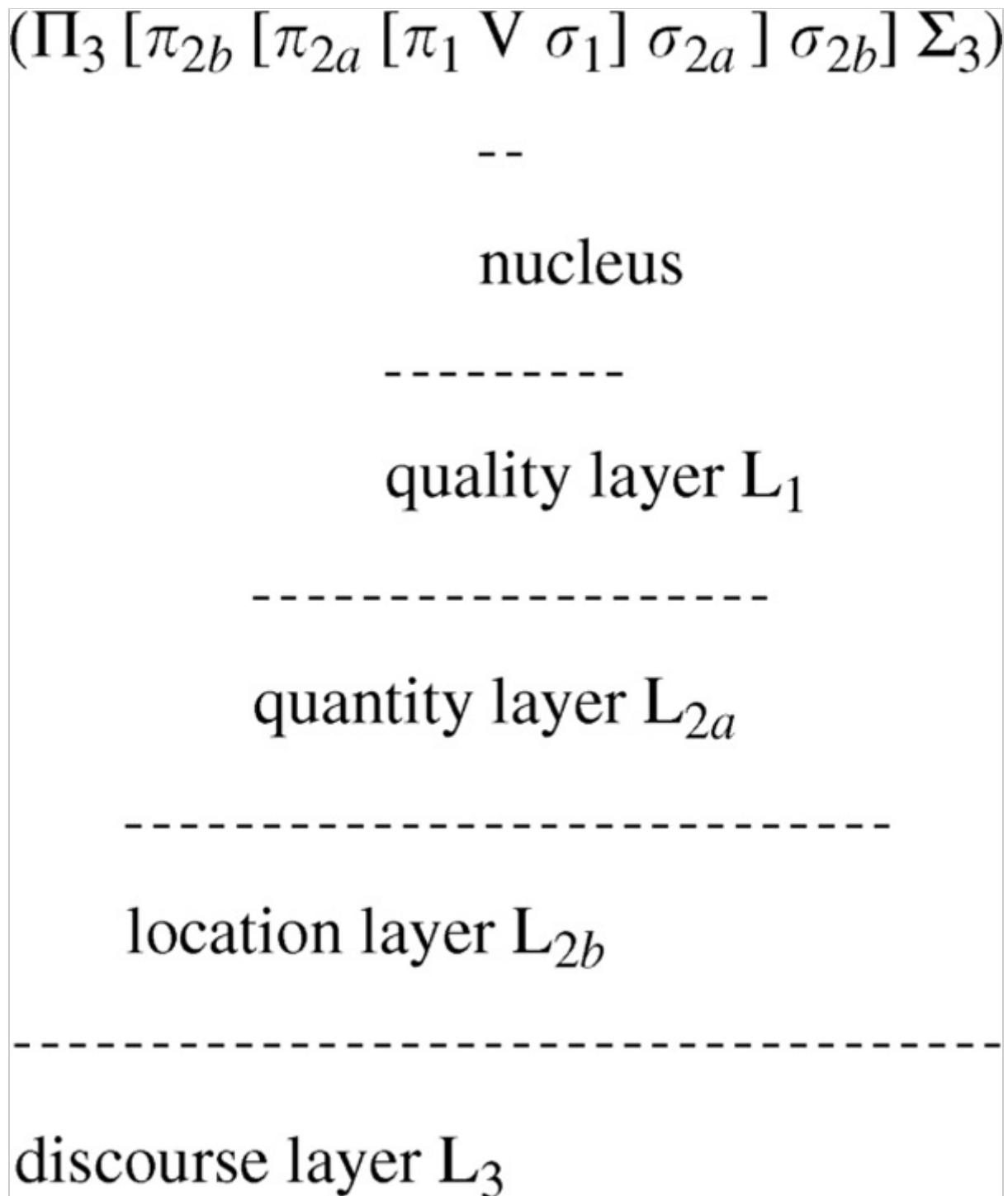


Fig. 11.3. The hierarchical structure of the clause (predication)

why ‘time is like space’ (Christensen 1993). This is not to deny that there are also some rather fundamental differences between time and space, or between temporal and spatial entities (see also Chapter 7, note 31). For example, whereas time is one-dimensional in that it extends in a linear fashion (between past and future), space is three-dimensional since it extends in all directions.

11.9. Noun phrase internal syntax

Chapters 8, 9, and 10 are concerned with the order of constituents in the noun phrase. Chapter 8 presents an overview of the main morpho-syntactic properties of the languages in the sample (such as the basic word order and the encoding of participants in the predicate) and a typology of constituent ordering principles. It also introduces the *Principle of Domain Integrity* (originally proposed in Rijkhoff 1990a), which in FG not

only accounts for syntactic phenomena that are captured by notions such as ‘constituency’ or ‘dependency’ in other grammatical theories, but also explains why the preferred position of embedded domains is in the periphery of the matrix domain. Finally, Chapter 8 discusses the notion of ‘iconicity’, since it is claimed that the *Principle of Domain Integrity* and the ordering principles that are discussed in Chapters 9 and 10 (*Principle of Head Proximity*, *Principle of Scope*) are both elaborations of a more general iconic principle that was already formulated by Behaghel (1932: 4).

Chapter 9 deals with the *Principle of Head Proximity*, which was originally formulated to account for some of Greenberg’s universals and the so-called Greenbergian word order correlations (Greenberg 1966a). The chapter focuses (p.345) on two predictions of the *Principle of Head Proximity*:

1. modifying adjectives occur immediately before or after the head noun (‘noun-adjective adjacency’) and
2. noun modifiers tend to precede the noun in a head-final (OV) language and to follow the noun in a head-initial (VO) language.

As to the first prediction, the only problematic language in the sample is Oromo (section 9.2.2.2).³ Although there is some evidence to suggest that the adjective is not a fully integrated noun modifier, too little is known about Oromo (and other languages in the same region with similar syntactic properties; section 9.2.3) to decide conclusively whether or not this language is a true counter-example to the first hypothesis.

It turned out that the second hypothesis, concerning the position of adnominal modifiers relative to the head noun, mostly holds for embedded modifiers such as adnominal possessor NPs and relative clauses, and not so much for modifiers of the simple NP, such as numerals and adjectives (see section 9.4 on demonstratives). As a possible explanation for the difference in syntactic behavior between embedded and non-embedded modifiers it was hypothesized that this is due to the fact that embedded modifiers are complex, referential structures, which are more difficult to process by the human parser than simple noun modifiers (cf. Dryer 1992; Hawkins 1994).

Additionally it was shown that the *Principle of Head Proximity* accounts for several of Dryer’s correlation pairs (Dryer 1992):

VERB patterner	OBJECT patterner
(a) copula verb	+ predicate (is + a teacher)
(b) ‘want’	+ VP (wants + to see Mary)
(c) tense/aspect aux. verb	+ VP (has + eaten dinner)
(d) negative auxiliary	+ VP (see Dryer 1992: 101)
(e) question particle	+ S (see Dryer 1992: 102–3)
(f) noun	+ genitive (father + of John)
(g) noun	+ relative clause (movies + that we saw)

These correlation pairs can be interpreted as a confirmation of the basic insight captured by the *Principle of Head Proximity*: that there is a significant cross-linguistic tendency to avoid the occurrence of linguistic material between heads of hierarchically ordered domains (see section 9.3.3.3).

Finally it was shown that Hawkins’s principle of *Early Immediate Constituents* (EIC) makes essentially the same predictions as the *Principle of Head Proximity* (section 9.3.4). Apart from the fact that both the principle of EIC and the *Principle of Head Proximity* account for the same set of correlation pairs (p.346) presented in Dryer (1992), the two principles also predict:

- a short-before-long ordering preference in VO-languages and a long-before-short preference in OV-languages;
- a strong tendency to avoid the occurrence of adnominal modifiers (notably the possessor NP and the relative clause) between the head of the clause and the head of the NP.

The third principle that determines ordering patterns in the noun phrase is the *Principle of Scope*, which predicts that modifiers occur next to the part of the expression they have in their scope. For the simple noun phrase this means that the actual ordering must be an instance of the general pattern [dem num ANA num dem], i.e. a pattern in which

- the localizing modifier (dem), which has the widest scope, is always the first or last in the sequence,
- the qualifying modifier (A), which only has the noun in its scope, is always adjacent to the noun (this is also predicted by the *Principle of Head Proximity*), and
- the quantifying modifier (num) never appears between A and N.

A number of potential counter-examples were investigated but in all cases (except one—cf. the section 9.2.3 on Oromo and other languages from the northeastern part of Africa) it could be demonstrated that the deviant order is due to the fact that the pattern includes a modifier that is not a morphologically free, syntactically integrated modifier of the simple NP (but e.g. a bound, apposed, or embedded modifier; see sections 9.2 and 10.2.4; see sections 10.2.1–2 on the position of the nominal aspect marker). Two predicted patterns are, however, not attested: [dem A N num] and [A N num dem]. These patterns indicate that a numeral may only appear after the noun if the adjective does so too (if [N num] order, then [N A] order), but no good explanation could be offered for this implication. Another instance of asymmetry was discussed in the two short sections on the internal syntax of complex NPs (sections 9.5 and 10.3), where it was found that heaviness appears to play a role in the syntax of possessor NPs and relative clauses. These embedded modifiers are attested in various positions in the NP when they precede the noun, but commonly appear after all non-embedded modifiers when they follow the noun.

11.10. Summing up

The most important findings of this study are:

1. Nouns are lexically specified for a particular *Seinsart* or ‘mode of being’ (just as verbs can be characterized in terms of *Aktionsart* or ‘mode of action’); section 2.6.
(p.347)
2. The *Seinsart* of noun can be defined in terms of the spatial features Shape and Homogeneity; section 2.4.
3. Within and across languages, four kinds of nouns (*Seinsarten*) are used when reference is made to a discrete spatial object: singular object nouns, set nouns, sort nouns, general nouns; section 2.4.
4. The special semantics of incorporated and predicate nouns is due to the fact that these nouns are not specified for a particular *Seinsart* (‘mode of being’); section 2.5.1.
5. There is a grammatical category of ‘nominal aspect’ (just as there is a grammatical category of ‘verbal aspect’); section 4.2.
6. Nominal aspect markers are used with set nouns (i.e. transnumeral nouns which can be in direct construction with a numeral) to indicate what *kind* of set is being referred to: a collective set (containing more than one member) or a singleton set (with only one member); section 4.2.1.2.
7. In languages that have regular instances of ‘number discord’ between the predicate and the argument(s), agreement is between the set entity (which is always singular) and the predicate, rather than between the individuals contained in the set and the predicate; section 4.2.1.2.
8. A language can only have a distinct class of adjectives, if it uses singular object nouns or set nouns to refer to discrete spatial objects (i.e. nouns that have the *Seinsart* feature + Shape); section 4.3.5.
9. If a language has flexible nouns rather than a rigid or distinct class of nouns, the nouns in that language are transnumeral and genderless (see also Hengeveld and Valstar forthcoming); section 2.2.4.
10. Modifiers in the NP can be divided over descriptive domains which correspond to three of Aristotle’s basic categories: the *ποιόν* ‘how it is’ = Quality, the *ποσόν* ‘how much/many it is’ = Quantity, and the *ποῦ* ‘where it is’ = Location. In addition there are modifiers which are essentially concerned with that it is’; these modifiers relate to the discourse status of the referent of the NP and indicate, for example, whether or not the referent of an NP already exists in the world of discourse; Chapter 7.
11. The NP can be analyzed in terms of four hierarchically organized (nested) layers around the head noun (the nucleus). The innermost layer is the quality layer, which is contained in the quantity layer, which in its turn is contained in the location layer. The outermost layer is concerned with discourse properties of the referent of the NP; Chapter 7.
12. Each layer accommodates grammatical and lexical manifestations of a particular modifier category (qualifying modifiers, quantifying modifiers, localizing modifiers, discourse modifiers); Chapter 7.

13. The underlying structure of the NP presented in this book is a *semantic* model of the NP, in which the nested layers capture the scope differences ([p.348](#)) between the various kinds of modifier categories. Modifiers represented in an outer layer have modifiers in the inner layer(s) in their scope; Chapter 7.

14. Up to a point (namely at the boundary between predication and proposition) the underlying structure of the sentence can be analyzed in a similar fashion to the NP, i.e. at some level of abstraction sentences and NPs have the same kind of layered organization, accommodating the same kind of modifier categories (relating to qualitative, quantitative, locative, and discourse properties of the referent of the NP or predication); Chapter 7.

15. The internal syntax of the simple, integral NP can be described by two principles: the *Principle of Domain Integrity*, which accounts for the fact that the noun and its modifiers (demonstrative, numeral, adjective) normally occur in an uninterrupted linear sequence, and the *Principle of Scope*, which predicts eight possible ordering patterns of demonstrative (dem), numeral (num), adjective (A), and head noun (N):

dem num A N	dem A N num	num A N dem	A N num dem
dem num N A	dem N A num	num N A dem	N A num dem

The fact that as a rule no other free modifier occurs between the (true) attributive adjective and the noun is also predicted by the *Principle of Head Proximity* (section 9.3.4.2.2). There is, however, no good explanation for the fact that the two italicized patterns [*dem A N num*, *A N num dem*] are unattested.

16. The *Principle of Head Proximity* predicts the preferred position of the relative clause and the possessor NP relative to the head noun: after the noun in a VO-language and before the noun in an OV-language. The position of non-embedded modifiers (demonstrative, numeral, adjective), however, does not seem to be affected by Head Proximity; for a possible explanation, see section 9.3.3.5.

17. The *Principle of Head Proximity* also predicts (among other things) a short-before-long ordering preference in VO-languages and a long-before-short preference in OV-languages (sections 9.1 and 9.3.4.2.2 and see Appendix to Chapter 9).

18. Complex NPs: Abstracting away from the OV/VO distinction (and from other ordering principles), the *Principle of Domain Integrity* predicts that the preferred position of an embedded domain (possessor NP, relative clause) is in the periphery of the matrix NP (i.e. not between constituents of the matrix NP). This holds true for a representative sample of European languages; section 9.5.

19. Complex NPs: Abstracting away from the OV/VO distinction (and from other ordering principles), the *Principle of Scope* predicts, however, that the demonstrative pronoun may precede or follow the embedded ([p.349](#)) localizing modifier (possessor NP, relative clause) if they occur on the same side of the head noun. Examples are given of all four possibilities; e.g. [dem Rel ... N] and [Rel dem ... N] are both attested in Turkish; [N ... dem Rel] is attested in the Celtic languages, and Nung has the order [N ... Rel dem]; section 10.3.

20. Complex NPs: Abstracting away from the OV/VO distinction (and from other ordering principles), the *Principle of Head Proximity* predicts that an embedded domain (possessor NP, relative clause) may appear between constituents of the simple (matrix) NP, in a position that gives a better ratio in terms of Head Proximity (so not in between the adjective and the head noun). Some evidence for this prediction was found in word order patterns from a representative sample of European languages, but only when the embedded modifier precedes the noun. When the embedded modifier follows the noun, the evidence suggests that there is a strong preference (possibly due to heaviness) for the embedded modifier to appear after all other modifiers in the complex NP; section 9.5.

Notes:

(1) Notice that this definition of nominal subclassification also covers nominal subcategories (count nouns, mass nouns, etc.) and the distinction between first and higher order nouns. However, nominal subcategories (*Seinsarten*) were discussed separately in Chapter 2. As to the classification of nouns into first and higher order nominals, I stated earlier that this book is restricted to nouns that are employed to refer to an entity whose ontological counterpart (or *Sein*-correlate) is a singular discrete spatial object.

(2) There seems to be a link between classifiers, noun classes (including genders), and number marking. As was mentioned in section 3.3, Greenberg (1972, 1978a) has outlined a possible diachronic path from (demonstrative plus) sortal classifier to gender marker, and before that he formulated a universal connecting gender and number (Greenberg 1966a: 95):

Universal 36. If a language has a category of gender, it always has the category of number.

Sortal classifiers typically occur with sort nouns, which are characterized by the feature – Shape and which are neutral regarding number distinctions. When the classifier combines with other constituents than the numeral, the demonstrative pronoun is the first modifier with which it occurs (Greenberg 1972). Over time the classified demonstrative may go through various stages of grammaticalization (Greenberg 1978a), becoming a nominal affix, the end result being a noun that is marked for class (or gender). Universal 36 suggest that the acquisition of gender goes hand in hand with the change from – Shape and + Shape (assuming that Greenberg interpreted aspect markers on set nouns as number marking). In short, it seems that the feature + Shape is transferred when a noun acquires gender.

- (3) But recall that Bender et al. (1976b: 144) state that in Oromo adjectives, numerals, and demonstratives all follow the noun in that order.

(p.350) References

Bibliography references:

- ABAEV, VASILII I. 1964. *A grammatical sketch of Ossetic* (*International Journal of American Linguistics*, 30/4). Bloomington: Indiana University.
- ABNEY, STEPHEN P. 1987. 'The English noun phrase in its sentential aspect'. Ph.D. dissertation, MIT.
- ADAMS, KAREN LEE. 1986. 'Numeral classifiers in Austroasiatic', in Craig 1986b: 241–62.
- 1989. *Systems of numeral classification in the Mon-Khmer, Nicobarese, and Aslian subfamilies of Austroasiatic* (Pacific Linguistics B-101). Canberra: Australian National University.
- and CONKLIN, NANCY FAIRES 1973. 'Toward a theory of natural classification', in Corum et al. 1973: 1–10.
- BECKER, ALTON L., and CONKLIN, NANCY F. 1975. 'Savoring the differences among classifiers'. Paper presented at the 8th International Conference on Sino-Tibetan Languages and Linguistics, Berkeley, Oct.: 1–19.
- AIKHENVALD, ALEXANDRA Y. 2000. *Classifiers: a typology of noun categorization devices*. Oxford: Oxford University Press.
- AL'KOR, J. P. (ed.). 1934. *Jazyki i pis'mennost narodov Severa III* (Languages and writings amongst the Northern people III). Leningrad: Instituta Narodov Severa (= Trudy po Lingvistike 3).
- ALLAN, KEITH. 1976. 'Collectivizing'. *Archivum Linguisticum*, 7: 99–117.
- 1977. 'Classifiers'. *Language*, 53/2: 285–311.
- ALLEN, BARBARA J., and GARDINER, DONNA B. 1977. 'Noun incorporation in Isleta'. *Work Papers of the Summer Institute of Linguistics* (University of North Dakota), 21: 45–69.
- FRANZ, DONALD G., GARDINER, DONNA B., and PERLMUTTER, DAVID M. 1990. 'Possessor ascension and syntactic levels in Southern Tiwa', in Postal and Joseph 1990: 321–83.
- AMEKA, FELIX. 1996. 'Body parts in Ewe grammar', in Chappell and McGregor 1996: 783–840.
- ANCEAUX, JOHANNES C. 1952. *The Wolio language: outline of grammatical description and texts*. The Hague: Martinus Nijhoff.
- ANDERSEN, HENNING. 1988. 'Center and periphery: adoption, diffusion and spread', in Fisiak 1988: 39–83.
- ANDERSON, JOHN M. 1971. *The grammar of case: towards a localistic theory*. London: Cambridge University Press.
- 1973. *An essay concerning aspect*. The Hague: Mouton.

- (ed.). 1982. *Language form and linguistic variation: papers dedicated to Angus McIntosh*. Amsterdam: Benjamins.
- ANDERSON, STEPHEN R. 1985. ‘Inflectional morphology’, in Shopen 1985: iii. 150–201.
- and KEENAN, EDWARD L. 1985. ‘Deixis’, in Shopen 1985: 259–308.
- ANDREWS, AVERY. 1985. ‘The major functions of the noun phrase’, in Shopen 1985: i. 62–154.
- (p.351) ANDRZEJEWSKI, BOGUMIŁ W. 1960. ‘The categories of number in noun forms in the Borana dialect of Galla’. *Africa*, 30: 62–75.
- ANWARD, JAN, MORAVCSIK, EDITH, and STASSEN, LEON. 1997. ‘Parts of speech: a challenge for typology’. *Linguistic Typology*, 1/2: 167–83.
- ARCHANGELI, DIANA, and LANGENDOEN, D. TERENCE (eds.). 1997. *Optimality Theory: an overview*. Malden, Mass.: Blackwell.
- ARISTAR, ANTHONY RODRIGUES. 1991. ‘On diachronic sources and synchronic pattern: an investigation into the origin of linguistic universals’. *Language*, 67/1: 1–33.
- ARNOLD, DOUG, ATKINSON, M., DURAND, J., GROVER, C., and SADLER, L. (eds.). 1989. *Essays on grammatical theory and universal grammar*. Oxford: Clarendon.
- ARNOTT, D. W. 1970. *The nominal and verbal systems of Fula*. Oxford: Oxford University Press.
- ARONOFF, MARK. 1994. *Morphology by itself: stems and inflectional classes*. London: MIT Press.
- ARONSON, HOWARD ISAAC (ed.). 1992. *The non-Slavic languages of the USSR*. Chicago: Chicago Linguistic Society.
- (ed.). 1994. *Non-Slavic languages of the USSR: papers from the fourth conference*. Columbus, Oh.: Slavica.
- ASHER, RON E. 1982. *Tamil* (Lingua Descriptive Studies 7). Amsterdam: North-Holland.
- and SIMPSON, J. M. Y. (eds.). 1994. *The encyclopedia of language and linguistics*, 10 vols. Oxford: Pergamon.
- AUSTERLITZ, ROBERT. 1980. ‘Typology and universals on a Eurasian east-west continuum’, in Brettschneider and Lehmann 1980: 235–414.
- BACH, EMMON. 1968. ‘Nouns and noun phrases’, in Bach and Harms 1968: 90–122.
- 1974. *Syntactic theory*. New York: Holt, Rinehart & Winston.
- and HARMS, ROBERT T. (eds.). 1968. *Universals in linguistic theory*. New York: Holt, Rinehart and Winston.
- JELINEK, ELOISE, KRATZER, ANGELIKA, and PARTEE, BARBARA (eds.). 1995. *Quantification in natural languages* (vols, i and ii of *Studies in Linguistics and Philosophy*). Dordrecht: Kluwer.
- BACKHOUSE, ANTHONY E. 1984. ‘Have all the adjectives gone?’. *Lingua*, 62: 169–86.

- BAHR, DONALD M. 1986. ‘Pima-Papago -ga, “alienability”’. *International Journal of American Linguistics*, 52/2: 161–71.
- BARKER, DIK. 1994. *Formal and computational aspects of functional grammar and language typology* (Studies in Language and Language Use 5). Amsterdam: IFOTT.
- and SIEWIERSKA, ANNA. 1992. ‘A database system for language typology’. *Working Paper 3 of Theme Group 2 (Constituent order) of the Eurotyp Programme in Language Typology*. Amsterdam: Instituut voor Algemene Taalwetenschap, Universiteit van Amsterdam.
- BARLOW, MICHAEL, and FERGUSON, CHARLES A. (eds.). 1988. *Agreement in natural languages: approaches, theories, descriptions*. Stanford, Calif.: CSLI.
- BARNES, JANET. 1990. ‘Classifiers in Tuyuca’, in Payne 1990: 273–92.
- BARRON, ROGER. 1980. ‘Das Phänomen klassifikatorischer Verben in nordamerikanischen Indianensprachen: ein typologischer Versuch’. *akup (=Arbeiten des Kölner Universalien-Projekts)* 38. Cologne: Institut für Sprachwissenschaft.
- (p.352) BARRON, ROGER and SERZISKO, FRITZ. 1982. ‘Noun classifiers in the Siouan languages’, in Seiler and Stachowiak 1982: 85–105.
- BARTSCH, RENATE. 1973. ‘The semantics and syntax of number and numbers’, in Kimball 1973b: 51–93.
- and VENNEMANN, THEO. 1972. *Semantic structures: a study in the relation between semantics and syntax*. Frankfurt am Main: Athenäum.
- BEALS, K., COOKE, G., KATHMAN, D., KITA, S., MCCULLOUGH, K., and TESTON, D. (eds.). 1993. *What we think, what we mean and how we say it: papers from the parasession on the correspondence of conceptual, semantic and grammatical representations*. Chicago: Chicago Linguistic Society.
- BECHERT, JOHANNES, BERNINI, GIULIANO, and BURIDANT, CLAUDE (eds.). 1990. *Toward a typology of European languages*. Berlin: Mouton de Gruyter.
- BECK, DAVID. 2000. ‘The syntax, semantics, and typology of adjectives in Upper Necaxa Totonac’. *Linguistic Typology*, 4/2: 213–50.
- BECKER, ALTON L. 1975. ‘A linguistic image of nature: the Burmese numerative classifier system’. *Linguistics*, 165: 109–21.
- BEEKES, ROBERT S. P. 1990. *Vergelijkende taalwetenschap: een inleiding in de vergelijkende Indo-europese Taalwetenschap*. Utrecht: Spectrum. [= *Comparative Indo-European linguistics: an introduction*. Amsterdam: Benjamins, 1996.]
- BEFFA, MARIE-LISE. 1982. ‘Présentation de la langue nivx’. *Études mongoles*, 13: 49–98.
- BEHAGHEL, OTTO. 1932. *Deutsche Syntax: eine geschichtliche Darstellung*. Vol. iv: *Wortstellung-Periodenbau*. Heidelberg: Carl Winter.
- BEHRENS, LEILA. 2000. *Typological parameters of genericity* (Arbeitspapier 37; NS). Cologne: Institut für Sprachwissenschaft, Universität zu Köln.

- BENDER, MARVIN L. (ed.). 1976. *The non-Semitic languages of Ethiopia* (Monograph 5, Occasional Papers Series). Committee on Ethiopian Studies, East Lansing-Carbondale, African Studies Center, Michigan State University.
- BOWEN, J. D., COOPER, R. L., and FERGUSON, C. A., (eds.). 1976a. *Language in Ethiopia*. London: Oxford University Press.
- ETEFFA, MULUGETA, and STINSON, D. LLOYD. 1976b. ‘Two Cushitic languages’, in Bender et al. 1976a: 130–54.
- BENNIS, HANS, and VAN LESSEN KLOEKE, W. U. S. (eds.). 1983. *Linguistics in the Netherlands 1983*. Dordrecht: Foris.
- BENTON, RICHARD A. 1968. ‘Numeral and attributive classifiers in Trukese’. *Oceanic Linguistics*, 7/2: 104–46.
- BENVENISTE, ÉMILE. 1966. ‘La Phrase nominale’, in Émile Benveniste, *Problèmes de linguistique générale*. Paris: Gallimard: 187–207.
- BENZING, JOHANNES. 1955. *Die tungusischen Sprachen: Versuch einer vergleichenden Grammatik*. Wiesbaden: Steiner.
- 1959. ‘Das Tschuwaschische’, in Deny et al. 1959: 695–751.
- 1985. *Kalmückische Grammatik zum Nachschlagen*. Wiesbaden: Harrassowitz.
- BERGER, HERMANN. 1998. *Die Burushaski-Sprache von Hunza und Nager*. Part 1: *Grammatik* (Neuindische Studien 13). Wiesbaden: Harrassowitz.
- BERMAN, HOWARD. 1972. ‘Relative clauses in Hittite’, in Peranteau et al. 1972: 1–8.
- BEUKEMA, FRITS, and COOPMANS, PETER (eds.). 1987. *Linguistics in the Netherlands 1987*. Dordrecht: Foris.
- (p.353) BEVER, THOMAS G. 1970. ‘The cognitive basis for linguistic structures’, in Hayes 1970: 279–362.
- BHAT, D. N. SHANKARA. 1994. *The adjectival category: criteria for differentiation and identification* (Studies in Language Companion Series 24). Amsterdam: Benjamins.
- BICKERTON, DEREK. 1981. *Roots of language*. Ann Arbor: Karoma.
- BIERMANN, ANNA. 1982. ‘Die grammatische Kategorie Numerus’, in Seiler and Lehmann 1982: 229–43.
- BIRD, CHARLES S. 1972. ‘The syntax and semantics of possession in Bambara’. Bloomington: Indiana University Linguistics Club.
- BIRK, D. B. W. 1976. *The MalakMalak language, Daly River (Western Arnhem Land)* (Pacific Linguistics B-45). Canberra: Australian National University.
- BISANG, WALTER. 1993. ‘Classifiers, quantifiers and class nouns in Hmong’. *Studies in Language*, 17/1: 1–51.
- 1996. ‘Areal typology and grammaticalization: processes of grammaticalization based on nouns and verbs in East and mainland South East Asian languages’. *Studies in Language*, 20/3: 519–97.

- 1999. ‘Classifiers in East and Southeast Asian languages: counting and beyond’, in Gvozdanović 1999: 113–85.
- BLAKE, BARRY J. 1983. ‘Structure and word order in Kalkatungu: the anatomy of a flat language’. *Australian Journal of Linguistics*, 3/2: 143–75.
- 1984. ‘Problems for possessor ascension: some Australian examples’. *Linguistics*, 22: 437–53.
- 1994. *Case*. Cambridge: Cambridge University Press.
- BLANKENSHIP, BARBARA. 1997. ‘Classificatory verbs in Cherokee’. *Anthropological Linguistics*, 39/1: 92–110.
- BLANSITT, EDWARD L., Jr. 1985. ‘SOVD languages’, in Elson 1985: 29–38.
- BOAS, FRANZ. 1911a. ‘Introduction’ to Boas 1911b: 5–83.
- (ed.). 1911b. *Handbook of American Indian languages*, part 1 (Bulletin 40 of the Smithsonian Institution). Washington: Bureau of American Ethnology. [Repr. Oosterhout: Anthropological Publications, 1969.]
- (ed.). 1922. *Handbook of American Indian languages*, part 2 (Bulletin 40 of the Smithsonian Institution). Washington: Bureau of American Ethnology. [Repr. Oosterhout: Anthropological Publications, 1969.]
- (ed.). 1941. *Handbook of American Indian languages*, part 4. New York: Augustin.
- 1947. ‘Kwakiutl grammar’. *Transactions of the American Philosophical Society*, NS 37/3.
- BODINE, ANN. 1975. ‘Sex differentiation in language’, in Thorne and Henley 1975:130–51.
- BOGORAS, WALDEMAR. 1922. ‘Chukchee’, in Boas 1922: 631–903.
- BOLINGER, DWIGHT. 1967. ‘Adjectives in English: attribution and predication’. *Lingua*, 18/1: 1–34.
- 1972. ‘Linear modification’, in Householder 1972: 31–50.
- BOLKESTEIN, A. MACHTELT, DE GROOT, C., and MACKENZIE, J. L. (eds.). 1985a. *Syntax and pragmatics in Functional Grammar* (Functional Grammar Series 1). Dordrecht: Foris.
- (eds.). 1985b. *Predicates and terms in Functional Grammar* (Functional Grammar Series 2). Dordrecht: Foris.
- (p.354) BONFANTE, GIULIANO. 1983. *The Etruscan language: an introduction*. Manchester: Manchester University Press.
- BONVILLAIN, NANCY 1973. *A grammar of Akwesasne Mohawk*. Ottawa: National Museum of Man.
- BOOIJ, GEERT, and VAN MARLE, JAAP (eds.) 1996. *Yearbook of morphology 1995*. Dordrecht: Kluwer.
- BORG, ALBERT. 1980. ‘An aspectual distinction in verbs and nouns in Maltese’. *Archivum Linguisticum*, 11/2: 107–17.

- BÖRJARS, KERSTI E. 1994. 'Feature distribution in Swedish noun phrases'. Ph.D. dissertation, University of Manchester.
- BOT BA Njock, HENRI MARCEL. 1977. 'L'Adjectif qualificatif dans trois langues bantu du nord-ouest: duálá, basaá, búlu', in Kotey and Der-Houssikian 1977: 207–25.
- BOUMA, SAAPKE. 1987. 'The discourse function of quantifier float: *tous* in French', in Beukema and Coopmans 1987: 31–40.
- BRADLEY, C. HENRY. 1970. *A linguistic sketch of Jicaltepec Mixtec*. Dallas: Summer Institute of Linguistics.
- BRAUNER, SIEGMUND. 1974. *Lehrbuch des Bambara*. Leipzig: Enzyklopädie.
- BRAUNMÜLLER, KURT. 1984. 'Morphologische Undurchsichtigkeit: ein Charakteristicum kleiner Sprachen', in Bra unmüller 1995: 48–68.
- 1995. *Beiträge zur Skandinavistischen Linguistik*. Oslo: Norus.
- BRETTSCHEIDER, GUNTER, and LEHMANN, CHRISTIAN (eds.). 1980. *Wege zur Universalienforschung: sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler*. Tübingen: Gunter Narr.
- BRIGHT, WILLIAM (ed.). 1992. *International Encyclopedia of Linguistics*, 4 vols. Oxford: Oxford University Press.
- BRINTON, LAUREL J. 1991. 'The mass/count distinction and Aktionsart: the grammar of iteratives and habituals'. *Belgian Journal of Linguistics*, 6: 47–69.
- BROSCHART, JÜRGEN. 1991. 'Noun, verb, and participation (a typology of Noun/Verb-distinction)', in Seiler and Prepper 1991: 65–137.
- 1997. 'Why Tongan does it differently: categorial distinctions in a language without nouns and verbs'. *Linguistic Typology*, 1: 123–65.
- BROWN, R. L., et al. (eds.). 1977. *Proceedings of the 1976 Mid-America Linguistics Conference*. Minneapolis: University of Minnesota Linguistics Department.
- BRUCE, LES. 1984. *The Alambak language of Papua New Guinea (East Sepik)* (Pacific Linguistics C-81). Canberra: The Australian National University.
- BULTEAU, R. 1953. *Cours d'Annamite* (4th edn.). Paris: Leroze.
- BUNT, HARRY. 1985. *Mass terms and model-theoretic semantics*. Cambridge: Cambridge University Press.
- BURKHARDT, HANS, SMITH, BARRY, et al. (eds.). 1991. *Handbook of metaphysics and ontology*. Munich: Philosophia Verlag.
- BURLING, ROBBINS. 1965. 'How to choose a Burmese numeral classifier', in Spiro 1965: 243–64.
- BYARUSHENG, E. R., DURANTI, A., and HYMAN, L. M. (eds.). 1977. 'Haya grammatical structure'. *Occasional Papers in Linguistics*, 6. Los Angeles: Dept. of Linguistics, University of Southern California.
- BYBEE, JOAN L. 1985. *Morphology: a study of the relation between meaning and form* (Typological Studies in Language 9). Amsterdam: Benjamins.

- (p.355) BYBEE, JOAN L. and DAHL, ÖSTEN. 1989. 'The creation of tense and aspect systems in the languages of the world'. *Studies in Language*, 13/1: 51–103.
- PAGLIUCA, WILLIAM, and PERKINS, REVERE D. 1990. 'On the asymmetries in the affixation of grammatical material', in Croft et al. 1990: 1–42.
- PERKINS, REVERE D., and PAGLIUCA, WILLIAM. 1994. *The evolution of grammar: tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- CAMPBELL, GEORGE L. 1991. *Compendium of the world's languages*. Vol. ii: *Maasai to Zuni*. London: Routledge.
- CAMPBELL, LYLE. 1985. *The Pipil language of El Salvador*. Berlin: Mouton de Gruyter.
- CAPELL, ARTHUR. 1965. 'A typology of concept domination'. *Lingua*, 15: 451–62.
- CARTER, RALPH M. 1976. 'Chipewyan classificatory verbs'. *International Journal of American Linguistics*, 42: 24–30.
- CASTRÉN, M. ALEXANDER. 1858. *Versuch einer jenissei-ostjakischen und kottischen Sprachlehre nebst Wörterverzeichnissen aus den genannten Sprachen* (Nordische Reisen und Forschungen 12). St Petersburg: Schieffner.
- CHAFFE, WALLACE L., and NICHOLS, JOHANNA (eds.). 1986. *Evidentiality: the linguistic coding of epistemology*. Norwood, NJ: Ablex.
- CHAO, YUEN REN. 1961. *Mandarin primer: an intensive course in spoken Chinese*. Cambridge, Mass.: Harvard University Press.
- CHAPPELL, HILARY, and MCGREGOR, WILLIAM (eds.). 1996. *The grammar of inalienability: a typological perspective on body part terms and the part–whole relation*. Berlin: Mouton de Gruyter.
- THOMPSON, SANDRA A. 1992. 'The semantics and pragmatics of associative *de* in Mandarin Chinese'. *Cahiers de linguistique Asie orientale*, 21/2: 199–229.
- CHILD, TUCKER. 1983. 'Noun class affix renewal in Southern West Atlantic', in Kaye et al. 1983: 17–29.
- 1995. *A grammar of Kisi: a Southern Atlantic language*. Berlin: Mouton de Gruyter.
- CHRISTENSEN, F. M. 1993. *Space-like time: consequences of, alternatives to, and arguments regarding the theory that time is like space*. Toronto: University of Toronto Press.
- CHRISTIE, J. J. 1970. 'Locative, possessive and existential in Swahili'. *Foundations of Language*, 6: 166–77.
- CHUNG, SANDRA, and TIMBERLAKE, ALAN. 1985. 'Tense, aspect, and mood', in Shopen 1985: iii. 202–58.
- CHURCHWARD, C. MAXWELL. 1941. *A new Fijian grammar*. Suva: Government Press. [Repr. 1973.]
- 1953. *Tongan grammar*. London: Oxford University Press.

- CHURCHWARD, S. 1951. *A Samoan grammar* (2nd edn., rev. and enlarged). Melbourne: Spectator Publishing Co. Pty. Ltd.
- CLARK, EVE V. 1970. ‘Locationals: a study of the relation between “existential”, “locative”, and “possessive” constructions’. *Working Papers in Language Universals*, 3 (Stanford University): L1–L36 + xiii.
- 1976. ‘Universal categories: on the semantics of classifiers and children’s early word meanings’, in Julland 1976: 449–62.
- (p.356) CLARK, EVE V. 1978. ‘Locationals: existential, locative, and possessive constructions’, in Greenberg et al. 1978: iv. 85–126.
- CLARK, HERBERT H., and CLARK, EVE V. 1977. *Psychology and language: an introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.
- CLAUDI, ULRIKE, and HEINE, BERND. 1986. ‘On the metaphorical basis of grammar’. *Studies in Language*, 10/2: 297–335.
- COLE, PETER (ed.). 1981. *Radical pragmatics*. New York: Academic Press.
- 1982. *Imbabura Quechua* (Lingua Descriptive Studies 5). Amsterdam: North-Holland.
- MORGAN, J. (eds.). 1975. *Speech acts* (Syntax and Semantics 3). New York: Academic Press.
- COMRIE, BERNARD. 1976. *Aspect*. Cambridge: Cambridge University Press.
- 1980. ‘Agreement, animacy, and voice’, in Brettschneider and Lehmann 1980: 229–34.
- 1981a. *The languages of the Soviet Union*. Cambridge: Cambridge University Press.
- 1981b. *Language universals and linguistic typology*. Oxford: Basil Blackwell.
- 1985. *Tense*. Cambridge: Cambridge University Press.
- (ed.). 1987. *The world’s major languages*. London: Croom Helm.
- 1988. ‘General features of the Uralic languages’, in Sinor 1988: 451–77.
- 1997. ‘Some problems in the theory and typology of numeral systems’, in Palek 1997: 41–56.
- and THOMPSON, SANDRA A. 1985. ‘Lexical nominalization’, in Shopen 1985: iii. 349–98.
- CONNOLLY, JOHN H. 1983. ‘Placement rules and syntactic templates’, in Dik 1983b: 247–66.
- and DIK, SIMON C. (eds.). 1989. *Functional Grammar and the computer* (Functional Grammar Series 10). Dordrecht: Foris.
- CONRAD, ROBERT J. (with WOGIGA, KEPAS). 1991. *An outline of Bukiyp grammar* (Pacific Linguistics C-113). Canberra: Australian National University.

- COOK, EUNG-DO. 1984. *A Sarcee grammar*. Vancouver: University of British Columbia Press.
- and WILHELM, ANDREA. 1998. ‘Noun incorporation: new evidence from Athapaskan’. *Studies in Language*, 22/21: 49–82.
- COPELAND, JAMES. 1998. ‘The grammaticalization of lexicalized manual gesture in Tarahumara’. Paper read at LSA/SSILA meeting, New York, 11 Jan.
- CORBETT, GREVILLE G. 1978. ‘Universals in the syntax of cardinal numbers’. *Lingua*, 46: 355–68.
- 1989. ‘An approach to the description of gender systems’, in Arnold et al. 1989: 53–89.
- 1991. *Gender*. Cambridge: Cambridge University Press.
- 2000. *Number*. Cambridge: Cambridge University Press.
- and FRASER, N. M. (eds.). 1993. *Heads in grammatical theory*. Cambridge: Cambridge University Press.
- and HAYWARD, RICHARD J. 1987. ‘Gender and number in Baiso’. *Lingua*, 72: 195–222.
- (p.357) CORNYN, WILLIAM STEWART, and ROOP, D. HAIGH. 1968. *Beginning Burmese*. New Haven: Yale University Press.
- CORUM, CLAUDIA, SMITH-STARK, T. CEDRIC, and WEISER, ANN (eds.). 1973. *You take the high node and I'll take the low node: papers from the 9th regional meeting of the Chicago Linguistic Society, April 13–15, 1973*. Chicago: Chicago Linguistic Society.
- COSTELLO, NANCY A. 1969. ‘The Katu noun phrase’. *Mon-Khmer Studies*, 3: 21–35.
- COURT, CHRISTOPHER. 1987. ‘Some classes of classifier in Iu Mien (Yao)’. *Linguistics of the Tibeto-Burman Area*, 10/2: 144–50.
- CRAIG, COLETTE GRINEVALD. 1977. *The structure of Jacaltec*. Austin: University of Texas Press.
- 1986a. ‘Jacaltec noun classifiers: a study in language and culture’, in Craig 1986b: 263–93.
- (ed.). 1986b. *Noun classes and categorization* (Typological Studies in Language 7). Amsterdam: Benjamins.
- 1992. ‘Classifiers in a functional perspective’, in Fortescue et al. 1992: 277–301.
- CREIDER, CHET A., and CREIDER, JANE TAPSUBEI. 1989. *A grammar of Nandi* (Nilo-Saharan 4). Hamburg: Buske.
- CREMERS, CRIT, and DEN DIKKEN, MARCEL (eds.). 1996. *Linguistics in the Netherlands 1996*. Amsterdam: Benjamins.
- CROCHETIÈRE, ANDRÉ, BOULANGER, JEAN-CLAUDE, and OUELLON, CONRAD (eds.). 1993. *Proceedings of the XVth International Congress of Linguists*. Sainte-Foy: Presses de l’Université Laval.

- CROFT, WILLIAM. 1988. 'Agreement vs. case marking and direct objects', in Barlow and Ferguson 1988: 159–79.
- 1990a. *Typology and universals*. Cambridge: Cambridge University Press.
- 1990b. 'A conceptual framework for grammatical categories (or: a taxonomy of propositional acts)'. *Journal of Semantics*, 1: 245–79.
- 1991. *Syntactic categories and grammatical relations: the cognitive organization of information*. Chicago: University of Chicago Press.
- 1995. 'Modern syntactic typology', in Shibatani and Bynon 1995: 85–145.
- 2000. 'Parts of speech as language universals and as language-particular categories', in Vogel and Comrie 2000: 65–102.
- DENNING, KEITH, and KEMMER, SUZANNE (eds.). 1990. *Studies in typology and diachrony: papers presented to Joseph H. Greenberg on his 75th birthday* (Typological Studies in Language 20). Amsterdam: Benjamins.
- CROWLEY, TERRY. 1985. 'Common noun phrase marking in Proto-Oceanic'. *Oceanic Linguistics*, 24/1–2: 135–93.
- CUSIC, DAVID. 1981. 'Verbal plurality and aspect'. Ph.D. dissertation, Stanford University.
- DAELEMAN, JAN. 1983. 'Tone-groups and tone-cases in a Bantu tone-language'. *LTL, Review of Applied Linguistics*, 60–1: 131–41.
- DAHL, ÖSTEN. 1975. 'On generics', in Keenan 1975: 99–111.
- 1985. *Tense and aspect systems*. Oxford: Blackwell.
- DAVIDSON, W., ELFORD, L. W., and HOIJER, H. 1963. 'Athapaskan classificatory verbs', in Hoijer et al. 1963: 102–9.
- DAVIES, JOHN. 1981. *Kobon* (Lingua Descriptive Studies 3). Amsterdam: North-Holland.
- (p.358) DEBOSE, CHARLES E. 1974. 'Papiamento plurals'. *Studies in African Linguistics*, Supplement 5: 67–73.
- DE GROOT, CASPER. 1989. *Predicate structure in a Functional Grammar of Hungarian* (Functional Grammar Series 11). Dordrecht: Foris.
- and LIMBURG, M. 1986. 'Pronominal elements: diachrony, typology, and formalization in Functional Grammar'. *Working Papers in Functional Grammar*, 12 (Universiteit van Amsterdam).
- and TOMMOLA, H. (eds.). 1984. *Aspect bound: a voyage into the realm of Germanic, Slavonic and Finno-Ugrian aspectology*. Dordrecht: Foris.
- DELANCEY, SCOTT. 1985. 'The analysis-synthesis-lexis cycle in Tibeto-Burman: a case study of motivated change', in Haiman 1985b: 367–89.
- DELSING, LARS-OLOF. 1993. 'The internal structure of noun phrases in the Scandinavian languages'. Doctoral dissertation, University of Lund.

- DEMUTH, KATHERINE, FARACLAS, NICHOLAS, and MARCHESE, LYNELL. 1986. ‘Niger-Congo noun class and agreement systems in language acquisition and historical change’, in Craig 1986b: 453–71.
- DENNING, KEITH, and KEMMER, SUZANNE (eds.). 1990. *On language: selected writings of Joseph H. Greenberg*. Stanford, Calif.: Stanford University Press.
- DENNY, J. PETER. 1976. ‘What are noun classifiers good for?’ *Papers from the 12th Meeting of the Chicago Linguistic Society*: 122–32.
- 1978. ‘Verb class meanings of the abstract finals in Ojibway inanimate intransitive verbs’. *International Journal of American Linguistics*, 44/4: 294–322.
- 1979. ‘The “extendedness” variable in classifier semantics: univerbal features and cultural variation’, in Mathiot 1979: 97–119.
- 1986. ‘The semantic role of noun classifiers’, in Craig 1986b: 297–308.
- and CREIDER, CHET A. 1986. ‘The semantics of noun classes in Proto-Bantu’, in Craig 1986b: 217–39.
- DENY, JEAN, GRÖNBECH, KAARE, SCHEEL, HELMUTH, and TOGAN, ZEKI VELIDI (eds.). 1959. *Philologiae Turcicae fundamenta*, vol. i. Wiesbaden: Steiner.
- DERBYSHIRE, DESMOND C. 1979. *Hixkaryana* (Lingua Descriptive Studies 1). Amsterdam: North-Holland.
- and PAYNE, DORIS L. 1990. ‘Noun classification systems of Amazonian languages’, in Payne 1990: 243–71.
- and PULLUM, GEOFFREY K. (eds.). 1986. *Handbook of Amazonian languages*, vol. i. Berlin: Mouton de Gruyter.
- (eds.). 1990. *Handbook of Amazonian languages*, vol. ii. Berlin: Mouton de Gruyter.
- (eds.). 1991. *Handbook of Amazonian languages*, vol. iii. Berlin: Mouton de Gruyter.
- (eds.). 1998. *Handbook of Amazonian languages*, vol. iv. Berlin: Mouton de Gruyter.
- DE REUSE, WILLEM J. 1994. ‘Noun incorporation’, in Asher and Simpson 1994: 2842–7.
- DEVITT, DANIEL C. 1994. ‘Copula constructions in cross-linguistic and diachronic perspective’. Ph.D. dissertation, State University of New York at Buffalo.
- DE VRIES, LOURENS. 1989. ‘Studies in Wambon and Kombai: aspects of two Papuan languages of Irian Jaya’. Ph.D. dissertation, University of Amsterdam.
- (p.359) DE VRIES, LOURENS. 1990. ‘Some remarks on direct quotation in Kombai’, in Pinkster and Genee 1990: 291–308.
- DIESSEL, HOLGER. 1999a. ‘The morphosyntax of demonstratives in synchrony and diachrony’. *Linguistic Typology*, 3/1: 1–49.

- 1999b. *Demonstratives: form, function, and grammaticalization*. Amsterdam: Benjamins.
- DIK, SIMON C. 1980a. ‘Non-verbal predicates’, in Simon C. Dik, *Studies in Functional Grammar*. New York: Academic Press: 90–112.
- 1980b. ‘From VSO to SVO’, in Simon C. Dik, *Studies in Functional Grammar*. New York: Academic Press: 152–77.
- 1983a. ‘Two constraints on relators and what they can do for us’, in Dik 1983b: 267–98.
- (ed.). 1983b. *Advances in Functional Grammar* (Publications in the Language Sciences 11). Dordrecht: Foris.
- 1986. ‘On the notion “functional explanation”’. *Belgian Journal of Linguistics*, 1: 11–52.
- 1989. *The theory of Functional Grammar*. Part 1: *The structure of the clause* (Functional Grammar Series 9). Dordrecht: Foris.
- 1997. *The theory of Functional Grammar* (2nd rev. edn., ed. Kees Hengeveld). Part 1: *The structure of the clause*. Part 2: *Complex and derived constructions*. Berlin: Mouton de Gruyter.
- and HENGEVELD, KEES. 1991. ‘The hierarchical structure of the clause and the typology of perception-verb complements’. *Linguistics*, 29: 231–59.
- DIMMENDAAL, GERRIT JAN. 1983. *The Turkana language* (Publications in African Languages and Linguistics 2). Dordrecht: Foris.
- DIXON, ROBERT M. W. (ed.). 1976. *Grammatical categories in Australian languages* (Linguistic Series 22). Canberra: Australian Institute of Aboriginal Studies.
- 1980. *The languages of Australia*. Cambridge: Cambridge University Press.
- 1982. *Where have all the adjectives gone? and other essays in semantics and syntax*. Berlin: Mouton de Gruyter.
- 1986. ‘Noun classes and noun classification in typological perspective’, in Craig 1986b: 105–12.
- 1988. *A grammar of Boumaa Fijian*. Chicago: University of Chicago Press.
- and AIKHENVALD, ALEXANDRA Y. (eds.). 1999. *The Amazonian languages* (Cambridge Language Surveys). Cambridge: Cambridge University Press.
- DMITRIJEV, N. K. 1939. *Stroj tureckogo jazyka* [Structure of the Turkic language]. Leningrad: Nauka.
- DONALDSON, T. 1980. *Ngiyambaa: the language of the Wangaaybuwan of New South Wales*. Cambridge: Cambridge University Press.
- DONNELLAN, KEITH. 1966. ‘Reference and definite descriptions’. *Philosophical Review*, 11: 281–304.
- DONNER, K. 1955–58. *Ketica: Materialien aus dem Ketischen oder Jenisseiostjakischen* (Supplement 1958). Helsinki: Suomalais-Ugrilainen Seura.

- DOTTER, FRANZ, and STRIEDER, MANFRED. 1988. ‘Zur Korrelation “Adposition–Position des Genitivs zum Kopfnomen”, *Papiere zur Linguistik*, 38/1: 51–60.
- (p.360) DOWNING, BRUCE T. 1978. ‘Some universals of relative clause structure’, in Greenberg et al. 1978: iv. 375–418.
- and FULLER, JUDITH. 1984. ‘Cultural contact and expansion of the Hmong lexicon’. *Linguistics club, University of Minnesota* 4/25/84.
- DOWTY, DAVID R., KARTTUNEN, LAURI, and ZWICKY, ARNOLD M. (eds.). 1985. *Natural language parsing: psychological, computational, and theoretical perspectives*. Cambridge: Cambridge University Press.
- DRABBE, PETER. 1959. *Grammar of the Asmat Language*, trans. Joseph Fichtner. Syracuse, NY: Our Lady of the Lake Press.
- DRESSLER, WOLFGANG. 1968. *Studien zur verbalen Pluralität: Iterativum, Distributivum, Durativum, Intensivum in der allgemeine Grammatik, im Lateinischen und Hethitischen*. Vienna: Österreichische Akademie der Wissenschaften.
- DREW, PAUL, and WOOTTON, A. (eds.). 1988. *Erving Goffman: exploring the interaction order*. Cambridge: Polity Press.
- DROHLA, WOLFGANG. 1949. ‘Die Kongruenz zwischen Nomen und Attribut sowie Subjekt und Prädikat in Hethitischen’. Dissertation, Philipps-Universität, Marburg.
- DROSSARD, WERNER. 1982. ‘Die Technik Masse/Messen auf der Dimension der Apprehension’, in Seiler and Lehmann 1982: 98–111.
- DRYER, MATTHEW S. 1988. ‘Object-verb order and adjective-noun order: dispelling a myth’. *Lingua*, 74 (special issue): 185–217.
- 1989a. ‘Article-noun order’, in Wiltshire et al. 1989a: 83–97.
- 1989b. ‘Plural words’. *Linguistics*, 27: 865–95.
- 1989c. ‘Discourse-governed word order and word order typology’. *Belgian Journal of Linguistics*, 4: 69–90.
- 1991. ‘SVO languages and the OV:VO typology’. *Journal of Linguistics*, 27: 443–82.
- 1992. ‘The Greenbergian word order correlations’. *Language*, 68/1: 81–138.
- 1997. ‘Aspects of word order in the languages of Europe’, in Siewierska 1997b: 283–319.
- 1998. ‘On the six-way word order typology’. *Studies in Language*, 21/1: 69–104.
- DUBOIS, JOHN W. 1985. ‘Competing motivations’, in Haiman 1985b: 343–65.
- 1987. ‘The discourse basis of ergativity’. *Language*, 63/4: 805–55.
- DU FEU, VERONICA M. 1987. ‘The determinants of the noun in Rapanui’. *Journal of the Polynesian Society*, 96: 473–95.
- 1989. ‘Verbal parameters expressed in the NP in Rapanui’. Paper read at the Colloquium on NP Structure, University of Manchester, Sept.
- 1996. *Rapanui*. London: Routledge.

- DUNN, MICHAEL. 2000. 'Chukchi women's language: a historical-comparative perspective'. *Anthropological Linguistics*, 42/3: 305–28.
- DURIE, MARK. 1986. 'The grammaticalization of number as a verbal category'. *Proceedings of the twelfth annual meeting of the Berkeley Linguistics Society*: 355–70.
- EDWARDS, ELIZABETH A. 1985. 'An animacy hierarchy in Haida'. *International Journal of American Linguistics*, 51: 398–400.
- ELLIOT, JENNIFER R. 2000. 'Realis and unrealis: forms and concepts of the grammaticalisation of reality'. *Linguistic Typology*, 4/1: 55–90.
- (p.361) ELSON, BENJAMIN F. (ed.). 1985. *Language in global perspective: papers in honor of the 50th anniversary of the Summer Institute of Linguistics 1935–1985*. Dallas: Summer Institute of Linguistics.
- EMENEAU, MURRAY B. 1951. *Studies in Vietnamese (Annamese) grammar*. Berkeley and Los Angeles: University of California Press.
- ENC, MÜRVET. 1991. 'The semantics of specificity'. *Linguistic Inquiry*, 22: 1–25.
- ENKVIST, NILS ERIK (ed.). 1982. *Impromptu speech: a symposium* (Publications of the Research Institute of the Åbo Akademi Foundation 78). Åbo: Åbo Akademi.
- ERBAUGH, MARY S. 1986. 'Taking stock: the development of Chinese noun classifiers historically and in young children', in Craig 1986b: 399–436.
- EVANS, NICHOLAS D. 1995. *A grammar of Kayardild: with historical-comparative notes on Tangkic*. Berlin: Mouton de Gruyter.
- 1996. 'The syntax and semantics of body part incorporation in Mayali', in Chappell and McGregor 1996: 65–109.
- EWERT, MANFRED, and HANSEN, FRED. 1993. 'On the linear order of the modifier-head-position in NPs', in Fanselow 1993: 161–81.
- FÄHNRICH, HEINZ. 1986. *Kurze Grammatik der georgischen Sprache*. Leipzig: Langenscheidt—Verlag Enzyklopädie.
- FANSELOW, GISBERT (ed.). 1993. *The parametrization of Universal Grammar* (Linguistik Aktuell 8). Amsterdam: Benjamins.
- FARSI, A. A. 1968. 'Classification of adjectives'. *Language Learning*, 18: 45–60.
- FERGUSON, CHARLES. 1976. 'The Ethiopian language area', in Bender et al. 1976a: 63–76.
- FISIAK, JACEK (ed.). 1988. *Historical dialectology, regional and social*. Berlin: Mouton de Gruyter.
- FLEISCHMAN, SUZANNE. 1989. 'Temporal distance: a basic linguistic metaphor'. *Studies in Language*, 13/1: 1–50.
- FODOR, ISTVÁN. 1959. 'The origin of grammatical gender (I–II)'. *Lingua*, 8: 1–41.
- FOLEY, WILLIAM A. 1980. 'Toward a universal typology of the noun phrase'. *Studies in Language*, 4/2: 171–99.

- 1986. *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.
- 1991. *The Yimas language of New Guinea*. Stanford, Calif.: Stanford University Press.
- (ed.). 1993. *The role of theory in language description*. Berlin: Mouton de Gruyter.
- 1997. *Anthropological linguistics: an introduction*. Oxford: Blackwell.
- FORTESCUE, MICHAEL. 1984. *West Greenlandic*. London: Croom Helm.
- 1992. ‘Aspect and superaspect in Koyukon: an application of the Functional Grammar model to a polysynthetic language’, in Fortescue et al. 1992: 99–141.
- HARDER, PETER, and KRISTOFFERSEN, LARS (eds.). 1992. *Layered structure and reference in Functional Grammar*. Amsterdam: Benjamins.
- FORTUNE, REO F. 1942. *Arapesh* (Publications of the American Ethnological Society 19). New York: Augustin.
- FOX, BARBARA. 1981. ‘Body part syntax: towards a universal characterization’. *Studies in Language*, 5/3: 323–42.
- (p.362) FOX, BARBARA. 1987. ‘The noun phrase accessibility hierarchy reinterpreted: subject primacy or the absolute hypothesis?’ *Language*, 63/4: 856–70.
- and THOMPSON, SANDRA A. 1990. ‘A discourse explanation of the grammar of relative clauses in English conversation’. *Language*, 66/2: 297–316.
- FRAJZYNGIER, ZYGMUNT. 1977. ‘The plural in Chadic’, in Newman and Ma Newman 1977: 37–56.
- 1985. ‘Ergativity, number, and agreement’. *Proceedings of the eleventh annual meeting of the Berkeley Linguistics Society*: 96–106.
- 1991. ‘The *de dicto* domain in language’, in Traugott and Heine 1991: 219–51.
- 1997. ‘Grammaticalization of number: from demonstratives to nominal and verbal plural’. *Linguistic Typology*, 1/2: 193–242.
- FRANK, PAUL. 1990. *Ika syntax* (Studies in the Languages of Columbia 1). Dallas: Summer Institute of Linguistics.
- FRAZIER, LYN. 1979. ‘Parsing and constraints on word order’. *University of Massachusetts Occasional Papers in Linguistics*, 5: 177–98.
- 1985. ‘Syntactic complexity’, in Dowty et al. 1985: 129–89.
- FREGE, GOTTLÖB. 1892. ‘Über Sinn und Bedeutung’. *Zeitschrift für Philosophic und philosophische Kritik*, 100: 25–50.
- FREI, HENRI. 1944. ‘Systèmes de déictiques’. *Acta Linguistica*, 4: 111–29.
- FRIEDRICH, JOHANNES. 1974. *Hethitisches Elementarbuch*. vol. i: *Kurzgefasste Grammatik* (3rd edn.). Heidelberg: Winter.
- FRIEDRICH, PAUL. 1970. ‘Shape in grammar’. *Language*, 46/2: 379–407.

- FRIES, PETER H. 1985. ‘Towards a discussion of the ordering of adjectives in the English noun phrase’, in Elson 1985: 123–33.
- GABLER, DARIUS. 1991. ‘Modi Essendi’, in Burkhardt et al. 1991: 567–8.
- GANDOUR, JACK, PETTY, SORANEE HOLASUIT, DARDARANANDA, ROCHANA, and DECHONGKIT, SUMALEE. 1984. ‘The acquisition of numeral classifiers in Thai’. *Linguistics*, 22: 455–79.
- GANGOPADHYAY, MALAYA. 1990. *The noun phrase in Bengali*. Delhi: Motilal Banarsi das.
- GAO, QIAN. 1994. ‘Chinese NP structure’. *Linguistics*, 32: 475–510.
- GARRETT, ANDREW. 1990. ‘The origin of NP split ergativity’. *Language*, 66/2: 261–95.
- GASPAROV, BORIS M. 1978. ‘Ustnaja reč’ kak semiotičeskij ob’ekt: Semantika nominacii i semiotika ustnoj reči’ (Spoken language as a semiotic object). *Acta et Commentationes Universitatis Tartuensis*, 442: Tartu 63–112.
- GEERTS, GUIDO, HAESERYN, W., DE ROOIJ, J., and VAN DEN TOORN, M. C. 1984. *Algemene Nederlandse Spraakkunst*. Groningen: Wolters-Noordhoff.
- GERNOT, WILHELM. 1995. ‘Suffixaufnahme in Hurrian and Urartian’, in Plank 1995: 113–35.
- GERSTNER-LINK, CLAUDIA. 1995. *Über Generizität: generische Nominalausdrücke in singulären Aussagen und generellen Aussagen*. Munich: Fink.
- . Forthcoming. ‘A typological approach to generics’. MS. University of Munich.
- GIL, DAVID. 1987. ‘Definiteness, noun phrase configurationality, and the count–mass distinction’, in Reuland and ter Meulen 1987: 254–69.
- GILDEA, SPIKE. 1993. ‘The development of tense markers from demonstrative pronouns in Panare (Cariban)’. *Studies in Language*, 17/1, 53–73.
- (p.363) GILLIGAN, GARY. 1987. ‘A cross-linguistic approach to the pro-drop parameter’. Ph.D. dissertation, USC.
- GIVÓN, TALMY. 1981a. ‘On the development of the numeral “one” as an indefinite marker’. *Folia Linguistica Historica*, 2/1: 35–55.
- . (ed.). 1981b. *Topic continuity in discourse*. Amsterdam: Benjamins.
- . 1984. *Syntax: a functional-typological introduction*, vol. i. Amsterdam: Benjamins.
- . 1988. ‘The pragmatics of word-order: predictability, importance and attention’, in Hammond et al. 1988: 243–84.
- . 1990. *Syntax: a functional-typological introduction*, vol. ii. Amsterdam: Benjamins.
- . 1991. ‘Serial verbs and the mental reality of “event”: grammatical vs. cognitive packaging’, in Traugott and Heine 1991: 81–127.

- GOODMAN, NELSON. 1966. *The structure of appearance* (2nd edn.). Indianapolis: Bobbs-Merrill.
- GORAL, DONALD R. 1978. ‘Numeral classifier systems: a Southeast Asian cross-linguistic analysis’. *Linguistics of the Tibeto-Burman Area*, 4/1: 1–72.
- GORBET, LARRY PAUL. 1976. *A grammar of Diegueño nominals*. New York: Garland.
- GORE, EDWARD C. 1926. *A Zande grammar*. London: Sheldon Press.
- GRAGG, GENE B. 1972. ‘Sumerian and selected Afro-Asiatic languages’, in Peranteau et al. 1972: 153–68.
- 1976. ‘Oromo of Wellegga’, in Bender 1976: 166–95.
- GREENBERG, JOSEPH H. 1966a. ‘Some universals of grammar with particular reference to the order of meaningful elements’, in Greenberg 1996b: 73–113. [Repr. Denning and Kemmer 1990: 40–70.]
- (ed.). 1966b. *Universals of language* (2nd edn.). Cambridge, Mass.: MIT.
- 1972. ‘Numeral classifiers and substantival number: problems in the genesis of a linguistic type’. *Working Papers on Language Universals* 9 (Stanford University), 1–39. [Repr. Denning and Kemmer 1990: 166–93.]
- 1974. ‘Studies in numeral systems, I: double numeral systems’. *Working Papers on Language Universals*, 14 (Stanford University): 75–89. [Repr. Denning and Kemmer 1990: 194–206.]
- 1975. ‘Dynamic aspects of word order in the numeral classifier’, in Li 1975: 29–45. [Repr. Denning and Kemmer 1990: 227–40.]
- 1977. ‘Niger-Congo noun class markers: prefixes, suffixes, both or neither’. *Studies in African Linguistics* (Supplement 7): 97–104.
- 1978a. ‘How does a language acquire gender markers?’, in Greenberg et al. 1978: iii. 48–82. [Repr. Denning and Kemmer 1990: 241–70.]
- 1978b. ‘Generalizations about numeral systems’, in Greenberg et al. 1978: iii. 250–95. [Repr. Denning and Kemmer 1990: 271–309.]
- 1981. ‘Nilo-Saharan moveable -k as a Stage III article (with a Penutian typological parallel)’. *Journal of African Languages and Linguistics*, 3: 105–12. [Repr. Denning and Kemmer 1990: 476–83.]
- 1985. ‘Some iconic relationships among place, time, and discourse deixis’, in Haiman 1985b: 271–87. [Repr. Denning and Kemmer 1990: 329–43.]
- (p.364) GREENBERG, JOSEPH H. 1989. ‘The internal and external syntax of numerical expressions: explaining language specific rules’. *Belgian Journal of Linguistics*, 4: 105–18.
- 1991a. ‘The last stages of grammatical elements; contrastive and expansive desemanticization’, in Traugott and Heine 1991: 301–14.
- 1991b. ‘The Semitic “intensive” as verbal plurality’, in Kaye 1991: 577–87.

- FERGUSON, CHARLES A., and MORAVCSIK, EDITH A. (eds.). 1978. *Universals of human language*. Vol. i: *Method and Theory*. Vol. ii: *Phonology*. Vol. iii: *Word structure*. Vol. iv: *Syntax*. Stanford, Calif.: Stanford University Press.
- GREGORES, EMMA, and SUÁREZ, JORGE A. 1967. *A description of colloquial Guarani*. The Hague: Mouton & Co.
- GRICE, H. PAUL. 1975. ‘Logic and conversation’, in Cole and Morgan 1975: 44–58.
- GRIERSON, GEORGE A. (ed.). 1906. *Linguistic survey of India*. Vol. iv: *Munda and Dravidian languages*. Delhi: Motilal Banarsi Dass: 185–9. [Repr. 1967, 1973.]
- GRIMES, BARBARA F. (ed.). 1988. *Ethnologue: languages of the world, plus ethnologue index* (11th edn.). Dallas: Summer Institute of Linguistics.
- GRINEVALD, COLETTE. 2000. ‘A morphosyntactic typology of classifiers’, in Senft 2000b: 50–92.
- GRÖNBECH, KAARE. 1936. *Die tiirkische Sprachbau I*. Copenhagen: Levin & Munksgaard.
- GROSJEAN, F., GROSJEAN, L., and LANE, H. 1979. ‘The patterns of silence: performance structures in sentence production’. *Cognitive Psychology*, 11: 58–81.
- GROSSMAN, ROBIN E., SAN, L. JAMES, and VANCE, TIMOTHY J. (eds.). 1975. *Papers from the eleventh regional meeting of the Chicago Linguistic Society, April 18–21, 1975*. Chicago: Chicago Linguistic Society.
- GRUZDEVA, EKATERINA. 1998. *Nivkh* (Languages of the World/Materials 111). Munich: LINCOM.
- GUNDEL, JEANETTE K., HOULIHAN, KATHLEEN, and SANDERS, GERALD. 1988. ‘On the function of marked and unmarked terms’, in Hammond et al. 1988: 285–301.
- GUY, J. B. M. 1974. *A grammar of the northern dialect of Sakao* (Pacific Linguistics B-33). Canberra: Australian National University.
- GVOZDANOVIĆ, JADRANKA (ed.). 1992. *Indo-European numerals*. Berlin: Mouton de Gruyter.
- (ed.). 1999. *Numeral types and changes worldwide* (Trends in Linguistics: Studies and Monographs 118). Berlin: Mouton de Gruyter.
- HAAS, MARY R. 1941. ‘Tunica’, in Boas 1941: 1–143.
- 1944. ‘Men’s and women’s speech in Koasati’. *Language*, 20: 142–9.
- 1948. ‘Classificatory verbs in Muskogee’. *International Journal of American Linguistics*, 14: 244–6.
- HAGÈGE, CLAUDE. 1974. ‘The “adjective” in some African languages’. *Studies in African Linguistics*, Supplement 5: 125–33.
- HAGMAN, ROY S. 1974. ‘Nama Hottentot grammar’. Ph.D. dissertation, Columbia University.
- HAIMAN, JOHN. 1980. ‘The iconicity of grammar: isomorphism and motivation’. *Language*, 56/3: 515–40.

- 1985a. *Natural syntax: iconicity and erosion*. Cambridge: Cambridge University Press.
- (p.365) HAIMAN, JOHN. 1985b. *Iconicity in syntax: proceedings of a symposium on iconicity in syntax, Stanford, June 24–6, 1983* (Typological Studies in Language 6). Amsterdam: Benjamins.
- 1988. ‘Rhaeto-Romance’, in Harris and Vincent 1988: 351–90.
- HALE, KENNETH L. 1975. ‘Gaps in grammar and culture’, in Kinkade et al. 1975: 295–315.
- 1976. ‘The adjoined relative clause in Australia’, in Dixon 1976: 78–105.
- 1983. ‘Warlpiri and the grammar of non-configurational languages’. *Natural Language and Linguistic Theory*, 1: 5–47.
- HALIM, AMRAN, CARRINGTON, LOIS, and WURM, S. A. (eds.). 1982. *Papers from the third International Conference on Austronesian Linguistics*. Vol. i: *Currents in Oceanic* (Pacific Linguistics C-74). Canberra: Australian National University.
- HAMMOND, MICHAEL, and NOONAN, MICHAEL (eds.). 1988. *Theoretical morphology: approaches in modern linguistics*. New York: Academic Press.
- MORAVCSIK, EDITH A., and WIRTH, JESSICA R. (eds.). 1988. *Studies in syntactic typology* (Typological Studies in Languages 17). Amsterdam: Benjamins.
- HANNAY, MIKE. 1985. ‘Inferrability, discourse-boundness, and sub-topics’, in Bolkestein et al. 1985a: 49–63.
- HARDER, PETER. 1996. *Functional semantics: a theory of meaning, structure and tense in English*. Berlin: Mouton de Gruyter.
- HARRIEHAUSEN, BETTINA. 1990. *Hmong Njua: syntaktische Analyse einer gesprochenen Sprache mithilfe datenverarbeitungstechnischer Mittel und sprachvergleichende Beschreibung des südostasiatischen Sprachraumes* (Linguistische Arbeiten 245). Tübingen: Niemeyer.
- HARRIS, ALICE CARMICHAEL. 1981. *Georgian syntax: a study in relational grammar* (Cambridge Studies in Linguistics 33). Cambridge: Cambridge University Press.
- 1985. *Diachronic syntax: the Kartvelian case* (Syntax and Semantics 18). New York: Academic Press.
- 1994. ‘On the history of relative clauses in Georgian’, in Aronson 1994: 130–42.
- HARRIS, MARTIN, and VINCENT, NIGEL (eds.). 1988. *The Romance languages*. London: Routledge.
- HARRISON, SHELDON P. 1982. ‘Proto-Oceanic *aki(ni) and the proto-Oceanic periphrastic causative’, in Halim et al. 1982: 179–230.
- 1988. ‘A plausible history for Micronesian possessive classifiers’. *Oceanic Linguistics*, 27/1–2: 63–78.
- HARVEY, MARC, and REID, NICHOLAS (eds.). 1997. *Nominal classification in aboriginal Australia* (Studies in Language Companion Series 37). Amsterdam:

Benjamins.

- HASHIMOTO, ANNE Y. 1971. 'Mandarin syntactic structures'. *Unicorn (Chi-lin)*, 8: 1–49.
- HASPELMATH, MARTIN. 1993. *A grammar of Lezgian* (Mouton Grammar Library 9). Berlin: Mouton de Gruyter.
- 1996. *From space to time: temporal adverbials in the world's languages* (LINCOM Studies in Theoretical Linguistics 03). Munich: Lincom Europa.
- HASPELMATH, MARTIN, KÖNIG, EKKEHARD, OESTERREICHER, W., and RAIBLE, W. (eds.). 2001. *Language typology and linguistic universals: an international handbook*, vol. i. Berlin: Walter de Gruyter.
- HAUGEN, EINAR I. 1976. *The Scandinavian languages*. London: Faber & Faber.
- (p.366) HAWKINS, EMILY A. 1982. *Pedagogical grammar of Hawaiian: recurrent problems*. Manoa: Hawaiian Studies Program, University of Hawaii.
- HAWKINS, JOHN A. 1983. *Word order universals: quantitative analyses of linguistic structure*. New York: Academic Press.
- 1988a. 'On explaining some left-right asymmetries in syntactic and morphological universals', in Hammond et al. 1988: 321–57.
- (ed.). 1988b. *Explaining language universals*. Oxford: Basil Blackwell.
- 1990. 'A parsing theory of word order universals'. *Linguistic Inquiry*, 21: 223–61.
- 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.
- 1997. 'Some issues in a performance theory of word order', in Siewierska 1997b: 729–81.
- HAYCOCK, B. G. 1978. 'The problem of the Meroitic language', in Thelwall 1978: 50–81.
- HAYES, JOHN R. (ed.). 1970. *Cognition and the development of language*. New York: John Wiley & Sons.
- HAYWARD, RICHARD J. 1984. *The Arbore language: a first investigation (including a vocabulary)*. Hamburg: Buske.
- and CORBETT, GREVILLE G. 1988. 'Resolution rules in Qafar'. *Linguistics*, 26/2: 259–79.
- HEATH, JEFFREY. 1984. *Functional grammar of Nunggubuyu*. Canberra: Australian Institute of Aboriginal Studies.
- HEIM, IRENE. 1982. 'The semantics of definite and indefinite noun phrases'. Ph.D. dissertation, University of Massachusetts at Amherst.
- HEINE, BERND. 1980. 'Determination in some East African languages', in Brettschneider and Lehmann 1980: 180–6.
- 1982. 'African noun class systems', in Seiler and Lehmann 1982: 189–216.

- 1990. ‘The dative in Ik and Kanuri’, in Croft et al. 1990: 129–49.
- 1997. *Cognitive foundations of grammar*. Oxford: Oxford University Press.
- SCHADEBERG, THILO, and WOLF, EKKEHARD (eds.). 1981. *Die Sprachen Afrikas*. Hamburg: Buske.
- CLAUDI, ULRIKE, and HÜNNEMEYER, FRIEDERIKE.
1991. *Grammaticalization*. Chicago: University of Chicago Press.
- HELD, WARREN. 1957. ‘The Hittite relative sentence’ (Language Dissertation 55). Baltimore: Waverly Press.
- HENGEVELD, KEE. 1986. ‘Copular verbs in a functional grammar of Spanish’. *Linguistics*, 24: 393–420.
- 1987. ‘Clause structure and modality in Functional Grammar’, in van der Auwera and Goossens 1987: 53–66.
- 1988. ‘Illocution, mood, and modality in a Functional Grammar of Spanish’. *Journal of Semantics*, 6/3: 227–69.
- 1989. ‘Layers and operators in Functional Grammar’. *Journal of Linguistics*, 25: 127–57.
- 1990a. ‘The hierarchical structure of utterances’, in Nuyts et al. 1990: 1–23.
- 1990b. ‘Semantic relations in non-verbal predictions’, in Nuyts et al. 1990: 101–22.
- 1992a. ‘Parts-of-speech’, in Fortescue et al. 1992: 29–55.
- (p.367) HENGEVELD, KEE. 1992b. *Non-verbal predication: theory, typology, diachrony*. Berlin: Mouton de Gruyter.
- 1998. ‘Terms and propositions’. Paper read at the eighth International Conference on Functional Grammar, Free University, Amsterdam, 8 July.
- and VALSTAR, MARIEKE. Forthcoming. ‘Parts of speech systems and lexical subclasses’. MS. Department of Linguistics, University of Amsterdam.
- RIJKHOFF, JAN, and SIEWIERSKA, ANNA, Forthcoming. ‘Parts-of-speech systems as a basic typological parameter’. Paper presented at the second meeting of the Association for Linguistic Typology (ALT II), Eugene, Oregon, 10–14 Sept. 1997.
- HERRING, SUSAN C. 1991. ‘Nominalization, relativization, and attribution in Lotha, Angami, and Burmese’. *Linguistics of the Tibeto-Burman Area*, 14/1: 55–72.
- HETZRON, ROBERT. 1967. ‘Agaw numerals and incongruence in Semitic’. *Journal of Semitic Studies*, 12/2: 169–97.
- 1978a. ‘The nominal system of Awngi (Southern Agaw)’. *Bulletin of the School of Oriental and African Studies*, 41/1: 121–41.
- 1978b. ‘On the relative order of adjectives’, in Seiler 1978b: 165–84.
- HEWITT, B. G. 1979. *Abkhaz* (Lingua Descriptive Studies 2). Amsterdam: North-Holland.

- HIMMELMANN, NIKOLAUS P. 1997. *Deiktikon, Artikel, Nominalphrase: zur Emergenz syntaktischer Struktur* (Linguistische Arbeiten 362). Tübingen: Niemeyer.
- 1998. ‘Regularity in irregularity: article use in adpositional phrases’. *Linguistic Typology*, 2/3: 315–53.
- HINDLE, DONALD. 1981. ‘A probabilistic grammar of noun phrases in spoken and written English’, in Sankoff and Cedergren 1981: 369–77.
- HJELMSLEV, LOUIS. 1935. *La Catégorie des cas*, Vol. i. Århus: Universitetsforlaget.
- HOCK, HANS HEINRICH. 1989. Review of S. B. Steever (1988), *The serial verb formation in the Dravidian languages*. *Language*, 65/2: 398–405.
- HOFFMANN, JOHANN. 1903. *Mundari grammar*. Calcutta: Secretariat Press.
- HOIJER, HARRY. 1945. ‘Classificatory verb stems in the Apachean languages’. *International Journal of American Linguistics*, 11: 13–23.
- et al. (eds.). 1963. *Studies in the Athapascan languages* (University of California Publications in Linguistics 29). Berkeley and Los Angeles: University of California Press.
- HOLM, JOHN. 1990. ‘Features in the noun phrase common to the Atlantic Creoles’. *Linguistics*, 28: 867–81.
- HOLMBERG, ANDERS, and RIJKHOFF, JAN. 1997. ‘Word order in the Germanic languages’, in Siewierska 1997b: 75–104.
- HOLMER, NILS MAGNUS. 1946. ‘Outline of a Cuna grammar’. *International Journal of American Linguistics*, 12: 185–97.
- 1947. *Critical and comparative grammar of the Cuna language*. Göteborg: Etnografiska Museet.
- HOLTON, GARY. 1999. ‘Categoriality of property words in a switch-adjective language’. *Linguistic Typology*, 3/3: 341–60.
- HONEY, P. J. 1972. ‘Word classes in Vietnamese’, in Householder 1972: 275–86.
- HOPE, EDWARD REGINALD. 1974. *The deep syntax of Lisu sentences: a transformational case grammar* (Pacific Linguistics B-34). Canberra: Linguistic Circle of Canberra.
- (p.368) HOPPER, PAUL (ed.). 1982. *Tense-aspect: between semantics and pragmatics* (Typological Studies in Language 1). Amsterdam: Benjamins.
- 1986. ‘Some discourse functions of classifiers in Malay’, in Craig 1986b: 309–25.
- and THOMPSON, SANDRA A. 1984. ‘The discourse basis for lexical categories in universal grammar’. *Language*, 60/3: 703–52.
- HOSKISON, JAMES T. 1983. ‘A grammar and dictionary of the Gude language’. Ph.D. dissertation, Ohio State University.
- HOUSEHOLDER, FRED WALTER (ed.). 1972. *Syntactic theory*. Vol. i: *Structuralist*. Harmondsworth: Penguin.

- HUANG, CHU-REN. 1989. *Mandarin Chinese de—a comparative study of current grammatical theories* (Academica Sinica Special Publications 93). Taipei: Institute of History and Philology, Academica Sinica.
- HUDSON, GROVER. 1976. ‘Highland East Cushitic’, in Bender 1976: 232–77.
- HUDSON, RICHARD A. 1976. ‘Beja’, in Bender 1976: 97–132.
- HUNDIUS, HARALD, and KÖLVER, ULRIKE. 1983. ‘Syntax and semantics of numeral classifiers in Thai’. *Studies in Language*, 7/2: 164–214.
- HURD, CONRAD. 1977. ‘Nasioi projectives’. *Oceanic Linguistics*, 16: 111–78.
- HURD, Ph. 1966. *Nasioi language course*. Port Moresby: Department of Information and Extension Services.
- HURFORD, JAMES R. 1987. *Language and number: the emergence of a cognitive system*. Oxford: Basil Blackwell.
- HYMAN, LARRY M. 1977. ‘Syntax of body parts’, in Byarushengo et al. 1977: 99–118.
- 1979. ‘Aghem grammatical structure’. *Southern California Occasional Papers in Linguistics*, 7 (UCLA).
- 1981. ‘Noni grammatical structure’. *Southern California Occasional Papers in Linguistics*, 9 (UCLA).
- HYMES, DELL (ed.). 1964. *Language in culture and society: a reader in linguistics and anthropology*. New York: Harper & Row.
- (ed.). 1971. *Pidginization and creolization of languages*. Cambridge: Cambridge University Press.
- IDE, SACHIKO, and HILL, BEVERLY (eds.). 1998. *Women’s languages in various parts of the world* (Special issue of *International Journal of the Sociology of Language*). Berlin: Mouton de Gruyter.
- ILJIC, ROBERT. 1994. ‘Quantification in Mandarin Chinese: two markers of plurality’. *Linguistics*, 32: 91–116.
- IMAI, M., and GENTNER, D. 1993. ‘Linguistic relativity vs. universal ontology: cross-linguistic studies of the object/substance distinction’, in Beals et al. 1993: 171–86.
- JACKENDOFF, RAY S. 1977. *X-bar syntax*. Cambridge, Mass.: MIT Press.
- 1990. *Semantic structures*. Cambridge, Mass.: MIT Press.
- JAHR, ERNST H. (ed.). 1992. *Language contact: theoretical and empirical studies*. Berlin: Mouton de Gruyter.
- JAISSE, ANNIE. 1987. ‘Hmong classifiers’. *Linguistics of the Tibeto-Burman Area*, 10/2: 169–76.
- JAKOBSON, ROMAN. 1971a. ‘Notes on Gilyak’, in Jakobson, 1971c: 79–97.
- 1971b. ‘Shifters, verbal categories, and the Russian verb’, in Jakobson 1971c: 130–47. [First part repr. Waugh and Monville-Burston 1990: 386–92.]

- (p.369) JAKOBSON, ROMAN. 1971c. *Selected writings*. Vol. ii: *Word and language*. The Hague: Mouton.
- JELINEK, ELOISE. 1995. ‘Quantification in Straits Salish’, in Bach et al. 1995: 487–540.
- and DEMERS, RICHARD A. 1994. ‘Predicates and pronominal arguments in Straits Salish’. *Language*, 70/4: 696–736.
- JENNER, PHILIP N., THOMPSON, LAURENCE C., and STAROSTA, STANLEY. 1976. *Austroasiatic studies*, part 2. Honolulu: University of Hawaii Press.
- JESPERSEN, OTTO. 1924. *The philosophy of grammar*. Chicago: University of Chicago Press.
- JOHANSON, LARS, and UTAS, BO (eds.). 1999. *Evidentials*. Berlin: Mouton de Gruyter.
- JONES, ROBERT B. 1970. ‘Classifier constructions in southeast Asia’. *Journal of the American Oriental Society*, 90: 1–12.
- JOSEPH, BRIAN D., and PHILIPPaki-WARBURTON, IRENE. 1987. *Modern Greek*. London: Croom Helm.
- JOSEPHS, LEWIS S. 1975. *Palauan reference grammar*. Honolulu: University of Hawaii Press.
- JUILLAND, ALPHONSE (ed.). 1976. *Linguistic studies offered to Joseph Greenberg on the occasion of his 60th birthday*. Saratoga, Calif.: Anima Libri.
- JURAFSKY, DANIEL. 1996. ‘Universal tendencies in the semantics of the diminutive’. *Language*, 72/3: 533–78.
- JUSTUS, C. 1976. ‘Topicalization and relativization in Hittite’, in Li 1976: 213–45.
- KAHR, J. C. 1976. ‘The renewal of case morphology: sources and constraints’. *Working Papers on Language Universals*, 20 (Stanford University): 107–51.
- KAHREL, PETER. 1985. ‘Indirect questions and relators’, in Bolkestein et al. 1985b: 165–81.
- KAMP, HANS, and REYLE, U. 1993. *From discourse to logic*. Dordrecht: Kluwer.
- KÄMPFE, HANS-RAINER, and VOLODIN, ALEXANDER P. 1995. *Abriß der tschuktschischen Grammatik auf der Basis der Schriftsprache* (Tunguso Sibirica 1). Wiesbaden: Harrassowitz.
- KANG, BEOM-MO. 1994. ‘Plurality and other semantic aspects of common nouns in Korean’, *Journal of East Asian Linguistics*, 3: 1–24.
- KARTTUNEN, LAURI. 1976. ‘Discourse referents’, in McCawley 1976: 363–87.
- KASTENHOLZ, RAIMUND. 1989. *Grundkurs Bambara (Manding) mit Texten*. Cologne: Köpfe.
- KAYE, ALAN S. (ed.). 1991. *Semitic studies in honor of Wolf Leslau on the occasion of his 85th birthday*. Wiesbaden: Harrassowitz.

- KAYE, J., KOOPMAN, H., SPORTICHE, D., and DUGAS, A. (eds.). 1983. *Current approaches to African linguistics*, vol. ii. Dordrecht: Foris.
- KAZENIN, KONSTANTIN I. 1994. ‘Order of modifiers in Godoberi noun phrase’, in Kibrik 1994: 144–50.
- KEENAN, EDWARD L. (ed.). 1975. *Formal semantics of natural language*. Cambridge: Cambridge University Press.
- 1979. ‘On surface form and logical form’. *Studies in the Linguistic Sciences*, 8: 1–41.
- 1985. ‘Relative clauses’, in Shopen 1985: ii. 141–70.
- and COMRIE, BERNARD. 1977. ‘Noun phrase accessibility and Universal Grammar’. *Linguistic Inquiry*, 8/1: 63–99.
- (p.370) KEENAN, EDWARD L. and COMRIE, BERNARD. 1979. ‘Data on noun phrase accessibility hierarchy’. *Language*, 55/2: 333–51.
- KEFER, MICHEL, and VANDER AUWERA, JOHAN (eds.). 1992. *Meaning and grammar: cross-linguistic perspectives*. Berlin: Mouton de Gruyter.
- KEIZER, M. EVELIEN. 1991. ‘Referring in Functional Grammar: how to define reference and referring expressions’. *Working Papers in Functional Grammar*, 43 (Universiteit van Amsterdam).
- 1992. ‘Predicates as referring expressions’, in Fortescue et al. 1992: 1–27.
- KIBRIK, ALEKSANDR E. (ed.). 1994. ‘Godoberi’s noun phrase structure’. *EUROTYPO Working Paper*, 7/21.
- KIEFER, FERENC. 1985. ‘The possessive in Hungarian: a problem for natural morphology’. *Acta Linguistica Academiae Scientiarum Hungaricae*, 35/1–2: 85–116.
- and KISS, KATALIN É (eds.). 1994. *The syntactic structure of Hungarian* (Syntax and Semantics 27). New York: Academic Press.
- KIM, NAM-KIL. 1987. ‘Korean’, in Comrie 1987: 811–98.
- KIM, YONG HWAN. 1993. ‘Dependent plural marking in Korean’. MA thesis, University of Texas at Austin.
- KIMBALL, GEOFFREY D. 1985. ‘A descriptive grammar of Koasati’. Ph.D. dissertation, Tulane University.
- 1987. ‘Men’s and women’s speech in Koasati: a reappraisal’. *International Journal of American Linguistics*, 53: 30–8.
- 1991. *Koasati grammar*. Lincoln: University of Nebraska Press.
- KIMBALL, JOHN P. 1973a. ‘Seven principles of surface structure parsing in natural language’. *Cognition*, 2: 15–47.
- (ed.). 1913b. *Syntax and semantics*, vol. ii. New York: Seminar.
- KINKADE, M. DALE. 1995. ‘A plethora of plurals: inflection for number in Upper Chehalis’. *Anthropological Linguistics*, 73/3: 347–65.

— et al. (eds.). 1975. *Linguistics and anthropology: in honour of C. F. Voegelin*. Lisse: Peter de Ridder.

KIRK, JOHN W. C. 1905. *A grammar of the Somali language*. Cambridge: Cambridge University Press.

KIRSNER, ROBERT S. 1985. ‘Iconicity and grammatical meaning’, in Haiman 1985b: 249–70.

KIYOMI, SETSUKO. 1995. ‘A new approach to reduplication: a semantic study of noun and verb reduplication in the Malayo-Polynesian languages’. *Linguistics*, 33: 1145–67.

KÖLVER, ULRIKE. 1978. ‘On Newari noun phrases’, in Seiler 1978b: 277–300.

— 1982. ‘Klassifikatorkonstruktionen in Thai, Vietnamesisch und Chinesisch: ein Beitrag zur Dimension der Apprehension’, in Seiler and Lehmann 1982: 160–85.

KOPTJEVSKAJA-TAMM, MARIA. 1993. *Nominalization*. London: Routledge.

— n.d. ‘The part-of-speech system in Chukchi’. MS. University of Stockholm.

KORNFILT, JAKLIN. 1987. ‘Turkish and the Turkic languages’, in Comrie 1987: 619–44.

KOTEY, PAUL F. A., and DER-HOUSSIKIAN, HAIG (eds.). 1977. *Language and linguistic problems in Africa: proceedings of the VII Conference on African Linguistics*. Columbia, SC: Hornbeam.

KOUWENBERG, SILVIA. 1991. ‘Berbice Dutch Creole: grammar, texts, and vocabulary’. Ph.D. dissertation, Universiteit van Amsterdam.

(p.371) KRÁMSKÝ, JIRI 1972. *The article and the concept of definiteness in language* (Janua Linguarum, Series Minor 125). The Hague: Mouton.

KRAUSS, MICHAEL E. 1968. ‘Noun classification systems in Athapascan, Eyal, Tlingit, and Haida verbs’. *International Journal of American Linguistics*, 34: 194–204.

KREJNOVIČ, E. A. 1934. ‘Nivxskij (giljackij) jazyk’ [The Nivjkh (Gilyak) language], in Al’kor 1934: 181–222.

KRIFKA, MANFRED. 1989. *Nominalreferenz und Zeitkonstitution: zur Semantik von Massentermen, Pluralterminen und Aspektklassen*. Munich: Fink.

KUH, HAKAN. 1986. ‘Plural copying in Korean’. *Harvard Studies in Korean Linguistics*, 2: 239–50.

KUHN, WILFRIED. 1982. ‘Formale Verfahren der Technik Kollektion’, in Seiler and Stachowiak 1982: 55–83.

KUIPER, FRANCISCUS B. J. 1962. *Nahali: a comparative study*. Amsterdam: Noord-Hollandsche Uitgevers Maatschappij.

KUNO, SUSUMO. 1974. ‘The position of relative clauses and conjunctions’. *Linguistic Inquiry*, 5: 117–36.

KURYŁOWICZ, JERZY. 1964. *The inflectional categories of Indo-European*. Heidelberg: Carl Winter.

— 1972. ‘The role of deictic elements in linguistic evolution’. *Semiotica*, 5: 174–83.

- 1976. ‘The plural in Semitic’, in Julland 1976: 95–102.
- KUTSCH LOJENGA, CONSTANCE. 1994. *Ngiti: a Central-Sudanic language of Zaire*. Cologne: Köppe.
- LABOV, WILLIAM. 1990. ‘On the adequacy of natural languages: I. The development of tense’, in Singler 1990: 1–58.
- LACKO, TIBOR. 1995. *The syntax of Hungarian noun phrases: a lexical-functional approach*. Frankfurt: Peter Lang.
- LA GALY, MICHAEL W., FOX, ROBERT A., and BRUCK, ANTHONY (eds.). 1974. *Papers from the tenth regional meeting of the Chicago Linguistic Society, April 19–21, 1974*. Chicago: Chicago Linguistic Society.
- LAKOFF, GEORGE. 1986. ‘Classifiers as a reflection of the mind’, in Craig 1986b: 13–51.
- 1987. *Women, fire and dangerous things: what categories reveal about the mind*. Chicago: University of Chicago Press.
- and JOHNSON, MARK. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.
- LAMBRECHT, KNUD. 1994. *Information structure and sentence form: topic, focus and the mental representations of discourse referents*. Cambridge: Cambridge University Press.
- LANDAR, H. 1964. ‘Seven Navajo verbs of eating’. *International Journal of American Linguistics*, 30: 94–6.
- LANDSBERG, MARGE E. (ed.). 1995. *Syntactic iconicity and linguistic freezes: the human dimension* (Studies in Anthropological Linguistics 9). Berlin: Mouton de Gruyter.
- LANG, ADRIENNE. 1975. *The semantics of classificatory verbs in Enga (and other Papua New Guinea languages)* (Pacific Linguistics B–39). Canberra: Australian National University.
- (p.372) LANGACKER, RONALD W. 1977. *Studies in Uto-Aztecan grammar*. Vol. i: *An overview of Uto-Aztecan grammar* (SIL Publications in Linguistics 56). Dallas: Summer Institute of Linguistics.
- 1987. ‘Nouns and verbs’. *Language*, 63/1: 53–94.
- 1990. ‘Cognitive grammar: the symbolic alternative’. *Studies in the Linguistic Sciences*, 20/2: 3–30.
- LAPPIN, SHALOM (ed.). 1988. *The syntax and semantics of NPs*. Special issue of *Lingua* (vol. 26).
- LAROCHE, EMMANUEL. 1982. ‘Épithètes et prédicats en hittite’, in Tischler 1982: 133–6.
- LAROCHE, J. 1958. *Grammaire des dialectes Mangbetu et Medje (suivie d'un manuel de conversation et d'un lexique)*. Tervuren: Koninklijk Museum van Belgisch-Congo.

- LASCARATOU, CHRYSSOULA. 1997. 'Basic characteristics of Modern Greek word order', in Siewierska 1997b: 151–71.
- LAURY, RITVA. 1997. *Demonstratives in interaction: the emergence of a definite article in Finnish* (Studies in Discourse and Grammar 7). Amsterdam: Benjamins.
- LÊ VĂN LÝ. 1960. *Le Parler vietnamien: essai d'une grammaire vietnamienne* (2nd edn.). Saigon: Institut des Recherches Historiques.
- LEE, HAN-GYU. 1991. 'Plural marker copying in Korean'. *Studies in the Linguistic Sciences*, 21/1: 81–105.
- LEE, HANSOL H. B. 1989. *Korean grammar*. Oxford: Oxford University Press.
- LEE, HYO SANH. 1993. 'The temporal system of noun-modifying (attributive) clauses in Korean from a typological perspective'. *Studies in Language*, 17/1: 75–110.
- LEECH, GEOFFREY N. 1969. *Towards a semantic description of English*. London: Longman.
- LEFEBVRE, CLAIRE, and MUYSKEN, PIETER. 1988. *Mixed categories: nominalization in Quechua*. Dordrecht: Kluwer.
- LEHMAN, FREDERICK K. 1990. 'Outline of a formal syntax of numeral expressions, with especial reference to the phenomenon of numeral classifiers'. *Linguistics of the Sino-Tibetan Area*, 13/1: 89–120.
- LEHMANN, CHRISTIAN. 1982a. 'Nominalisierung—Typisierung von Propositionen', in Seiler and Lehmann 1982: 64–82.
- 1982b. 'Universal and typological aspects of agreement', in Seiler and Stachowiak 1982: 201–67.
- 1984. *Der Relativsatz: Typologie seiner Strukturen, Theorie seiner Funktionen, Kompendium seiner Grammatik*. Tübingen: Narr.
- 1986. 'On the typology of relative clauses'. *Linguistics*, 24: 663–80.
- 1995. *Thoughts on grammaticalization*. Munich: Lincom Europa.
- LEHMANN, WINFRED P. 1973. 'A structural principle and its implications'. *Language*, 49/1: 42–66.
- 1978a. 'The great underlying ground-plans', in Lehmann 1978b: 3–55.
- (ed.). 1978b. *Syntactic typology*. Austin: University of Texas Press.
- LEINONEN, MARJA. 1985. 'Finnish and Russian as they are spoken: from linguistic to cultural typology'. *Scando-Slavica*, 31: 117–44.
- LEVINSON, STEPHEN C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- 1988. 'Putting linguistics on a proper footing: explorations in Goffman's concepts of participation', in Drew and Woottton 1988: 161–227.
- (p.373) LEVINSON, STEPHEN C. 1992. 'Primer for the field investigation of spatial description and conception'. *Pragmatics*, 2/1: 5–47.

- LEWIS, GEOFFREY L. 1967. *Turkish grammar*. Oxford: Oxford University Press. [Repr. 1984.]
- LI, CHARLES N. (ed.). 1975. *Word order and word order change*. Austin: University of Texas Press.
- (ed.). 1976. *Subject and topic*. New York: Academic Press.
- and THOMPSON, SANDRA A. 1989. *Mandarin Chinese: a functional reference grammar*. Berkeley and Los Angeles: University of California Press. [1st pub. 1981.]
- LICHTENBERK, FRANTISEK. 1983. ‘Relational classifiers’. *Lingua*, 60: 147–76.
- LIMBURG, MACHIEL. 1983. ‘The parameter preposition/postposition in word order typology’, in Bennis and van Lessen Kloek 1983: 149–59.
- 1985. ‘On the notion “relator” and the expression of the genitive relation’, in Bolkestein et al. 1985b: 147–63.
- LINDSLEY, J. R. 1975. ‘Producing simple utterances: how far ahead do we plan?’ *Cognitive Psychology*, 1: 1–19.
- LINELL, PER. 1982. *The written language bias in linguistics* (Studies in Communication (= SIC) 2). Linköping: Department of Communication Studies.
- LJUNG, MAGNUS. 1974. ‘Some remarks on antonymy’. *Language*, 50/1: 74–88.
- LÖBEL, ELISABETH. 1999. ‘Classifiers versus noun classes: a case study in Vietnamese’, in Unterbeck and Rissanen 1999: 259–319.
- LÖNNQVIST, BARBARA. 1982. ‘The study of Russian impromptu speech in Soviet linguistics’, in Enkvist 1982: 341–53.
- LORIMER, DAVID L. R. 1935–8. *The Burushaski language*, 3 vols. Cambridge, Mass.: Harvard University Press.
- LOTZ, JÁNOS. 1939. *Das ungarische Sprachsystem*. Stockholm: University of Stockholm. [Repr. Bloomington, Ind.: Eurolingua, 1988.]
- LUCY, JOHN A. 1992. *Grammatical categories and cognition: a case study of the linguistic relativity hypothesis* (Studies in the Social and Cultural Foundations of Language 13). Cambridge: Cambridge University Press.
- 2000. ‘Systems of nominal classification: a concluding discussion’, in Senft 2000b: 326–41.
- LUKAS, JOHANNES. 1928. ‘Transition und Intransition in Kanuri’. *Wiener Zeitschrift für die Kunde des Morgenlandes*, 35: 213–14.
- LUONG, HY V. 1987. ‘Plural markers and personal pronouns in Vietnamese personal reference: an analysis of pragmatic ambiguity and native models’. *Anthropological Linguistics*, 29/1: 49–70.
- LURAGHI, SILVIA. 1990. *Old Hittite sentence structure*. London: Routledge.
- LYONS, CHRISTOPHER. 1986. ‘The syntax of English genitive constructions’. *Linguistics*, 22: 123–43.
- 1999. *Definiteness*. Cambridge: Cambridge University Press.

- LYONS, JOHN. 1967. 'A note on possessive, existential and locative sentences'. *Foundations of Language*, 3: 390–6.
- 1977. *Semantics*, 2 vols. Cambridge: Cambridge University Press.
- MCCAWLEY, JAMES D. 1968. 'The role of semantics in a grammar', in Bach and Harms 1968: 124–69.
- (p.374) MCCAWLEY, JAMES D. 1971. 'Where do noun phrases come from?', in Steinberg and Jakobovits 1971: 217–31.
- (ed.). 1976. *Notes from the linguistic underground* (Syntax and Semantics 7). New York: Academic Press.
- 1992. 'Justifying part-of-speech assignments in Mandarin Chinese'. *Journal of Chinese Linguistics*, 20/2: 211–45.
- MCCOOL, G. J. 1984. 'The Old French demonstrative system: a semantic analysis'. *Word*, 35: 205–17.
- 1993. 'The French demonstrative system: from Old to Modern French'. *Word*, 44: 31–40.
- MACDONALD, LORNA. 1990. *A grammar of Tauya* (Mouton Grammar Library 6). Berlin: Mouton de Gruyter.
- MCGREGOR, WILLIAM B. 1989. 'Phrase fracturing in Gooniyandi', in Marácz and Muysken 1989: 207–22.
- 1997a. *Semiotic grammar*. Oxford: Oxford University Press.
- 1997b. 'Functions of noun phrase discontinuity in Gooniyandi'. *Functions of Language*, 4/1: 83–114.
- MACKENZIE, J. LACHLAN. 1983. 'Nominal predicates in a Functional Grammar of English', in Dik 1983b: 31–51.
- 1987a. 'Nominalization and basic constituent ordering', in van der Auwera and Goossens 1987: 93–105.
- 1987b. 'The representation of nominal predicates in the fund'. *Working Papers in Functional Grammar*, 25 (Universiteit van Amsterdam).
- 1992a. 'Places and things', in Fortescue et al. 1992: 253–76.
- 1992b. 'English spatial prepositions in Functional Grammar'. *Working Papers in Functional Grammar*, 46 (Universiteit van Amsterdam).
- MALCHUKOV, ANDREJ L. 2000. *Dependency reversal in noun-attribute constructions: towards a typology* (LINCOM Studies in Language Typology 03). Munich: Lincom Europa.
- MALLINSON, GRAHAM. 1986. 'Languages with and without extraposition'. *Folia Linguistica*, 20/1–2: 147–63.
- and BLAKE, BARRY J. 1981. *Language typology*. Amsterdam: North Holland.
- MANZELLI, GIANGUIDO. 1990. 'Possessive adnominal modifiers', in Bechert et al. 1990: 63–111.

- MARÁCZ, LÁSZLÓ, and MUYSKEN, PIETER (eds.). 1989. *Configurationality: the typology of asymmetries* (Studies in Generative Grammar 34). Foris: Dordrecht.
- MARTIN, LAURA. 1998. ‘Irrealis constructions in Mocho (Mayan)’. *Anthropological Linguistics*, 40/2 (special issue: ‘Symposium or Irrealis’): 198–213.
- MARTIN, SAMUEL E., and LEE, YOUNG-SOOK. 1969. *Beginning Korean*. New Haven: Yale University Press.
- MATHIOT, MADELEINE. 1979. *Ethnolinguistics: Boas, Sapir and Whorf revisited*. The Hague: Mouton.
- MATRAS, YARON (ed.). 1998. *The Romani element in non-standard speech* (Sondersprachenforschung 3). Wiesbaden: Harrassowitz.
- MATTHEWS, GEORGE HUBERT. 1965. *Hidatsa syntax*. The Hague: Mouton.
- MATTHEWS, PETER H. 1981. *Syntax*. Cambridge: Cambridge University Press.
- (p.375) MATTISSEN, JOHANNA, and DROSSARD, WERNER. 1998. *Lexical and syntactic categories in Nivkh (Gilyak)* (Theorie des Lexikons: Arbeiten des Sonderforschungsbereich 282, no. 85). Düsseldorf: Heinrich Heine Universität.
- MAXWELL, DANIEL N. 1979. ‘Strategies of relativization and NP accessibility’. *Language*, 55/2: 352–71.
- 1982. ‘Implications of NP accessibility for diachronic syntax’. *Folia Linguistica Historica*, 3/2: 135–52.
- MEILLET, ANTOINE. 1906. ‘La Phrase nominale en indo-européen’. *Mémoires de la Société de Linguistique de Paris*, 14: 1–26.
- 1964. ‘The feminine gender in the Indo-European languages’, in Hymes 1964: 124.
- 1967. *The comparative method in historical linguistics*, trans. Gordon B. Ford, Jr. Paris: Champion.
- MEINHOF, CARL. 1909. *Lehrbuch der Nama-Sprache*. Berlin: Reiner.
- 1912. *Die Sprachen der Hamiten*. Hamburg: Hamburger Kolonialinstitut.
- MEISTERFELD, REINHARD. 1998. *Numerus und Nominalaspekt: eine Studie zur romanischen Apprehension*. Tübingen: Niemeyer.
- MENGES, KARL HEINRICH. 1968. *The Turkic languages and peoples: an introduction to Turkic studies*. Wiesbaden: Harrassowitz.
- MENNINGER, KARL. 1969. *Number words and number symbols: a cultural history of numbers*, trans. Paul Broneer. Cambridge, Mass.: MIT Press.
- MERLAN, FRANCESCA. 1982. *Mangarayi* (Lingua Descriptive Studies 4). Amsterdam: North-Holland.
- 1983. *Ngalakan grammar, texts, and vocabulary* (Pacific Linguistics B-89). Canberra: Australian National University.
- ROBERTS, STEVEN POWELL, and RUMSEY, ALAN. 1997. ‘New Guinea “classificatory verbs” and Australian noun classification: a typological comparison’, in Harvey and Reid 1997: 63–103.

- Miao Language Team. 1972. ‘A brief description of the Miao language’, in Purnell 1972: 1–25.
- MILLER, CAROLYN P. 1964. ‘The substantive phrase in Brôu’. *Mon-Khmer Studies*, 1: 63–80.
- MITHUN, MARIANNE. 1984. ‘The evolution of noun incorporation’. *Language*, 60: 847–94.
- 1986. ‘The convergence of noun classification systems’, in Craig 1986b: 379–97.
- 1988. ‘Lexical categories and the evolution of number marking’, in Hammond and Noonan 1988: 211–34.
- 1992. ‘Is basic word order universal?’, in Payne 1992: 15–61. [Also in Tomlin 1987: 281–328.]
- 1996. ‘Multiple reflections of alienability in Mohawk’, in Chappell and McGregor 1996: 633–49.
- 1999. *The languages of Native North America* (Cambridge Language Surveys). Cambridge: Cambridge University Press.
- 2000. ‘Noun and verb in Iroquoian languages: multicategorisation from multiple criteria’, in Vogel and Comrie 2000: 397–420.
- MITHUN WILLIAMS, MARIANNE. 1976. *A grammar of Tuscarora*. New York: Garland.
- (p.376) MOORE, T. E. 1972. ‘Speeded recognition of ungrammaticality’. *Journal of Verbal Learning and Verbal Behavior*, 11: 550–60.
- MORAVCSIK, EDITH A. 1969. ‘Determination’. *Working Papers on Language Universals*, 1: 64–130.
- 1978a. ‘Reduplicative constructions’, in Greenberg et al. 1978: iii. 297–334.
- 1978b. ‘Agreement’, in Greenberg et al. 1978: iv. 331–74.
- 1994. ‘Inflectional morphology in the Hungarian noun phrase: a typological assessment’. *EuroTyp Working Papers*, 7/22: 1–58 (in each article pages are numbered separately).
- 1997. ‘Parts and wholes in the Hungarian noun phrase: a typological study’, in Palek 1997: 307–24.
- MORENO, MARTINO M. 1939. *Grammatica teorico-pratica della lingua galla, con esercizi*. Milan: n.p. [Repr. Ridgewood, NJ, 1964.]
- MORETON, REBECCA LARCHE. 1999. ‘The specifics of Specificity’. *LINGUIST List*, 10/12, 5 Jan.
- MOSEL, ULRIKE, and HOVDHAUGEN, EVEN. 1992. *Samoan reference grammar*. Oslo: Universitetsforlaget AS.
- MOSER, MARY B. 1977. ‘Articles in Seri’. *Work Papers of the Summer Institute of Linguistics, University of North Dakota*, 21: 1–22.

- MOURELATOS, ALEXANDER P. D. 1981. ‘Events, processes, and states’, in Tedeschi and Zaenen 1981: 191–212.
- MOUS, MAARTEN. 1993. *A grammar of Iraqw* (Cushitic Language Studies 9). Hamburg: Buske.
- MÜHLHÄUSLER, PETER. 1996. *Linguistic ecology; language change and linguistic imperialism in the Pacific region*. London: Routledge.
- NEDERGAARD THOMSEN, OLE. 1992. ‘Unit accentuation as an expression device for predicate formation in Danish’, in Fortescue et al. 1992: 173–229.
- NEWMAN, PAUL. 1990. *Nominal and verbal plurality in Chadic* (Publications in African Languages and Linguistics 12). Dordrecht: Foris.
- 1991. ‘An interview with Joseph Greenberg’. *California Linguistic Newsletter*, 23/1: 50–60.
- and MA NEWMAN, ROXANA (eds.). 1977. *Papers in Chadic linguistics*, Leiden: Afrika-Studiecentrum.
- NEY, J. W. 1982. ‘The order of adjectives and adverbs in English’. *Forum Linguisticum*, 6: 217–57.
- 1983. ‘Optionality and choice in the selection of order of adjectives in English’. *General Linguistics*, 23: 94–128.
- NGUYỄN ĐĂNG LIÊM. 1969. *Vietnamese grammar: a combined tagmemic and transformational approach. A contrastive analysis of English and Vietnamese*, vol. ii (Pacific Linguistics C-4). Canberra: Australian National University. [Repr. 1975.]
- NGUYỄN ĐÌNH HOÀ. 1987. ‘Vietnamese’, in Comrie 1987: 777–96.
- 1997. *Vietnamese* (London Oriental and African Language Library 9). Amsterdam: John Benjamins.
- NICHOLS, JOHANNA. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.
- NOONAN, MICHAEL P. 1985. ‘Complementation’, in Shopen 1985: ii. 42–140.
- (p.377) NOONAN, MICHAEL P. 1992. *A grammar of Lango* (Mouton Grammar Library 7). Berlin: Mouton de Gruyter.
- NORMAN, JERRY. 1988. *Chinese*. Cambridge: Cambridge University Press.
- NUYTS, JAN. 1992. *Aspects of a cognitive-pragmatic theory of language*. Amsterdam: Benjamins.
- and DE SCHUTTER, GEORGES (eds.). 1987. *Getting one’s words into line: on word order in Functional Grammar* (Functional Grammar Series 5). Dordrecht: Foris.
- BOLKESTEIN, A. MACHTELT, and VET, CO (eds.). 1990. *Layers and levels of representation in language theory: a functional view*. Amsterdam: Benjamins.
- OCHS, ELINOR. 1987. ‘The impact of stratification and socialization on men’s and women’s speech in Western Samoa’, in Philips et al. 1987: 50–70.

- O'DOWD, ELIZABETH. 1992. 'The syntactic metaphor of subordination: a typological study'. *Lingua*, 86: 47–80.
- OKELL, JOHN. 1969. *A reference grammar of Colloquial Burmese*, 2 vols. London: Oxford University Press.
- O'NEIL, W. 1982. 'Simplifying the grammar of English', in Anderson 1982: 285–306.
- OSGOOD, CHARLES (ed.). 1946. *Linguistic structures of native America* (Viking Fund Publications in Anthropology 6). New York: Viking.
- OWENS, JONATHAN. 1980. 'Observations on tone in the Booran dialect of Oromo (Galla)'. *African Language Studies*, 17: 141–96.
- 1985. *A grammar of Harar Oromo (Northeastern Ethiopia)* (Cushitic Language Studies 4). Hamburg: Buske.
- PALEK, BOHUMIL (ed.). 1995. *Item order in natural languages: proceedings of LP '94*. Prague: Charles University Press.
- (ed.). 1997. *Typology: prototypes, item orderings and universals: proceedings of LP '96*. Prague: Charles University Press.
- PALMER, FRANK R. 1986. *Mood and modality*. Cambridge: Cambridge University Press.
- 1992. Review of William Croft (1991), *Syntactic categories and grammatical relations*. *Journal of Linguistics*, 28: 269–73.
- PANFILOV, V. S. 1962–5. *Grammatika nivxskogo jazyka* [Nivkh grammar], 2 vols. Moscow: Mauk.
- PARKER, STEVE. 1998. 'Notes on determiners in Chamicuro'. *Notes on Linguistics*, 80: 8–22.
- 1999. 'On the behavior of definite articles in Chamicuro'. *Language*, 75/3: 552–62.
- PAWLEY, A. 1993. 'A language which defies description by ordinary means', in Foley 1993: 87–130.
- PAYNE, DORIS L. 1986. 'Noun classification in Yagua', in Craig 1986b: 113–31.
- 1987. 'Noun classification in the Western Amazon'. *Language Sciences*, 9: 21–44 (special issue): *Comparative linguistics of South American Indian languages*, ed. Mary Ritchie Key.
- (ed.). 1990. *Amazonian linguistics: studies in lowland South American languages*. Austin: University of Texas Press.
- (ed.). 1992. *Pragmatics of word order flexibility*. Amsterdam: Benjamins.
- PAYNE, JOHN. 1994. 'Nouns and noun phrases', in Asher and Simpson 1994: v. 2848–55.
- (p.378) PERANTEAU, PAUL M., LEVI, JUDITH N., and PHARES, GLORIA C. (eds.). 1972. *The Chicago wh- hunt: papers from the relative clause festival. A paravolume to Papers from the eighth regional meeting*. Chicago: Chicago Linguistic Society.

- PERKINS, REVERE D. 1992. *Deixis, grammar, and culture* (Typological Studies in Language 24). Amsterdam: Benjamins.
- PHILIPS, SUSAN U., STEELE, SUSAN, and TANZ, CHRISTINE (eds.).
1987. *Language, gender, and sex in a comparative perspective* (Studies in the Social and Cultural Foundations of Language 4). Cambridge: Cambridge University Press.
- PIAU, JULIE ANN. 1981–2. ‘Kuman classificatory verbs’. *Language and Linguistics in Melanesia* (formerly *Kivung*), 13/1–2: 3–31.
- PINKSTER, HARM, and GENEE, INGE (eds.). 1990. *Unity in diversity: papers presented to Simon C. Dik on his 50th birthday*. Dordrecht: Foris.
- PLANK, FRANS. 1979. ‘The functional basis of case systems and declension classes: from Latin to Old French’. *Linguistics*, 17/7–8: 611–640.
- 1984. ‘24 Grundsätzliche Bemerkungen zur Wortarten-Frage’. *Leuvense Bijdragen*, 73/4: 489–520.
- 1990. ‘Suffix copying as a mirror-image phenomenon’. *Linguistics*, 28: 1039–45.
- 1991. ‘On determiners: 1. Ellipsis and inflections, 2. Co-occurrence of possessives’. *EUROTYPO Working Paper*, 7/11.
- (ed.). 1995. *Double case: agreement by Sufftxaufnahme*. Oxford: Oxford University Press.
- 1997. ‘Word classes in typology: recommended reading’. *Linguistic Typology*, 1/2: 185–92.
- (ed.). Forthcoming. *Noun phrase structure in the languages of Europe*. Berlin: Mouton de Gruyter.
- POPPE, NICHOLAS. 1964. ‘Der altaische Sprachtyp’, in Spuler 1964: 1–16.
- POSTAL, P. M., and JOSEPH, B. D. (eds.). 1990. *Studies in Relational Grammar 3*. Chicago: University of Chicago Press.
- PRASITHRATHSINT, AMARA. 2000. ‘Adjectives as verbs in Thai’. *Linguistic Typology*, 4/2: 251–71.
- PRINCE, ELLEN F. 1981. ‘Toward a taxonomy of Given-New information’, in Cole 1981: 223–55.
- PURNELL, HERBERT C., Jr. (ed.). 1972. *Miao and Yao linguistic studies* (Selected articles in Chinese, trans. Chang Yu-hung and Chu Kwo-ray. Data Paper Number 88). Ithaca, NY: Department of Asian Studies, Cornell University.
- PUSTET, REGINA. 1989. *Die Morphosyntax des ‘Adjektivs’ im Sprachvergleich*. Frankfurt am Main: Lang.
- 2000. ‘How arbitrary is lexical categorization? Verbs vs. adjectives’. *Linguistic Typology*, 4/2: 175–212.
- QUESADA, J. DIEGO. 1999. ‘Chibchan, with special reference to participant-highlighting’. *Linguistic Typology*, 3/2: 209–58.

- QUINE, WILLARD VAN ORMAN. 1960. *Word and object*. Cambridge, Mass.: MIT Press.
- 1969. *Ontological relativity and other essays*. New York: Columbia University Press.
- QUIRK, RANDOLPH, GREENBAUM, SIDNEY, LEECH, GEOFFREY, and SVARTVIK, JAN. 1985. *A comprehensive grammar of the English language*. London: Longman.
- (p.379) RABEL-HEYMANN, LILI 1977. ‘Gender in Khasi nouns’. *Mon-Khmer Studies*, 6: 247–72.
- RAMSTEDT, GUSTAV J. 1939. *A Korean grammar* (Mémoires de la Société Finno-Ougrienne 82). Helsinki: Suomalais-Ugrilainen Seura.
- RANKIN, ROBERT L. 1977. ‘From verb to auxiliary, to noun classifier and definite article: grammaticalization of the Siouan verbs Sit, Stand, Lie’, in Brown et al. 1977: 273–83.
- RAPAPORT, WILLIAM J. 1985/6. ‘Non-existent objects and epistemological ontology’. *Grazer Philosophische Studien*, 25–6: 61–95.
- RATLIFF, MARTHA. 1987. ‘Tone sandhi compounding in White Hmong’. *Linguistics of the Tibeto-Burman Area*, 10/2: 71–105.
- 1991. ‘Cov, the underspecified noun, and syntactic flexibility in Hmong’. *Journal of the American Oriental Society*, 111/4: 694–703.
- RAUH, GISA (ed.). 1983. *Essays on deixis*. Tübingen: Gunter Narr.
- RAUSCH, P. J. 1912. ‘Die Sprache von Südost-Bougainville, Deutsche Salominsinseln’. *Anthropos*, 7: 105–34, 585–616, 964–94.
- REBUSCHI, G. 1990. Review of Mario Saltarelli (1988), *Basque. Lingua*, 80: 79–94.
- REESINK, GER P. 1984. ‘Structures and their functions in Usan, a Papuan language of Papua New Guinea’. Ph.D. dissertation, University of Amsterdam.
- REH, MECHTHILD. 1983. ‘Krongo: a VSO language with postpositions’. *Journal of African languages and Linguistics*, 5: 45–55.
- 1985. *Die Krongo-Sprache (Nîinò Mó-di): Beschreibung, Texte, Wörterverzeichniss*. Berlin: Reimer.
- 1994. ‘A grammatical sketch of Deiga’. *Afrika und Übersee* 77: 197–261.
- REULAND, ERIC J., and TER MEULEN, ALICE (eds.). 1987. *The representation of (in)definiteness*. Cambridge, Mass.: MIT Press.
- RIJKHOFF, JAN. 1984. ‘The principle of Head Proximity’. MA thesis, University of Amsterdam.
- 1986. ‘The principle of Head Proximity’. *Belgian Journal of Linguistics*, 1: 95–125 (= *Working Papers in Functional Grammar*, 14, Universiteit van Amsterdam).
- 1987. ‘Word order tendencies in two Prefield subtypes’, in Nuysts and de Schutter 1987: 1–15.

- 1988. ‘A typology of operators’. *Working Papers in Functional Grammar*, 29, Universiteit van Amsterdam.
- 1989. ‘The identification of referents’, in Connolly and Dik 1989: 229–46.
- 1990a. ‘Explaining word order in the noun phrase’. *Linguistics*, 28: 5–42.
- 1990b. ‘Toward a unified analysis of terms and predications’, in Nuyts et al. 1990: 165–91.
- 1991. ‘Nominal aspect’. *Journal of Semantics*, 8: 291–309.
- 1993. “Number” disagreement’, in Crochetière et al. 1993: 274–6.
- 1995. ‘De ontbrekende Categorie: over symmetric in the onderliggende struktuur’ (The missing category: on symmetry in the underlying structure). *Gramma/TTT-tijdschrift voor taalwetenschap*, 4/3: 199–222.
- 1997. ‘Order in the noun phrase of the languages of Europe’, in Siewierska 1997b: 321–82.
- 1998. ‘Bystander deixis’, in Matras 1998: 51–67.
- 2000. ‘When can a language have adjectives? An implicational universal’, in Vogel and Comrie 2000: 217–57.
- (p.380) RIJKHOFF, JAN. 2001. ‘Dimensions of adnominal modification’, in Haspelmath et al. 2001: 522–33.
- and BARKER, DIK. 1998. ‘Language sampling’. *Linguistic Typology*, 2–3: 263–314.
- — — HENGEVELD, KEEs, and KAHREL, PETER. 1993. ‘A method of language sampling’. *Studies in Language*, 17/1: 169–203.
- RIJKSBARON, ALBERT. 1989. *Aristotle, verb meaning and functional grammar*. Amsterdam: Gieben.
- RISCHEL, JØRGEN. 1995. *Minor Mlabri: a hunter-gatherer language of Northern Indo-china*. Copenhagen: Museum Tuscalanum.
- RISSELADA, RODIE. 1984. ‘Coordination and juxtaposition of adjectives in the Latin NP’ *Glotta*, 62/3–4: 202–31.
- ROSS, JOHN R. 1973. ‘The penthouse principle and the order of constituents’, in Corum et al. 1973: 397–422.
- ROSS, MALCOLM. 1996. ‘Adjectives with possessor nouns’. *LINGUIST List*: 7/1678: Subject 7.1678.
- ROWLANDS, EVAN C. 1959. *A grammar of Gambian Mandinka*. London: School of Oriental and African Studies. [2nd impr. 1969.]
- RUHLEN, MERRITT. 1987. *A guide to the world’s languages*. Vol. i: *Classification*. London: Arnold.
- 1991. *A guide to the world’s languages*. Vol. i: *Classification (with a postscript on recent developments)*. London: Arnold.

- SACKMANN, ROBIN. 1996a. ‘The problem of “adjectives” in Mandarin Chinese’, in Sackmann 1996b: 257–75.
- (ed.). 1996b. *Theoretical linguistics and grammatical description: papers in honour of Hans-Heinrich Lieb*. Amsterdam: Benjamins.
- SALTARELLI, MARIO. 1988. *Basque*. London: Croom Helm.
- SANCHES, MARY, and SLOBIN, LINDA. 1973. ‘Numeral classifiers and plural marking: an implicational universal’. *Working Papers in Language Universals*, 11 (Stanford University): 1–22.
- SANDERS, GERALD. 1976. ‘Functional constraints in grammar’, in Julland 1976: 161–78.
- SANDS, KRISTINA. 1995. ‘Nominal classification in Australia’. *Anthropological Linguistics*, 37: 247–346.
- SANKOFF, DAVID, and CEDERGREN, HENRIETTA (eds.). 1981. *Variation omnibus* (Current Inquiry into Language, Linguistics and Human Communication 40). Edmonton: Linguistic Research.
- SAPIR, EDWARD. 1921. *Language: an introduction to the study of speech*. New York: Harcourt, Brace & World. [Repr. London: Granada, 1982.]
- 1922. ‘The Takelma language of southwestern Oregon’, in Boas 1922: 1–296.
- and SWADESH, MORRIS. 1946. ‘American Indian grammatical categories’. *Word*, 2/2: 103–12.
- SASSE, HANS-JÜRGEN. 1976. ‘Dasenech’, in Bender 1976: 1196–221.
- 1981. ‘Die kushitischen Sprachen’, in Heine et al. 1981: 187–215.
- 1988. *Der irokesische Sprachtyp* (Arbeitspapier 9; NS). Cologne: Institut für Sprachwissenschaft, Universität zu Köln.
- 1993. ‘Das Nomen—eine universale Kategorie?’ *Sprachtypologie und Universalienforschung (STUF)*, 46/3: 187–221.
- (p.381) SASSE, HANS-JÜRGEN. 2001. *Recent activity in the theory of aspect: accomplishments, achievements, or just non-progressive state?* (Arbeitspapier 40; NS). Cologne: Institut für Sprachwissenschaft, Universität zu Köln.
- SAUL, JANICE E., and FREIBERGER WILSON, NANCY. 1980. *Nung grammar* (Summer Institute of Linguistics Publications in Linguistics 62). Dallas: Summer Institute of Linguistics.
- SCHACHTER, PAUL. 1985. ‘Parts-of-speech systems’, in Shopen 1985: i. 3–61.
- SCHADEBERG, THILO C. 1981. ‘The classification of the Kadugli language group’, in Schadeberg and Bender 1981: 291–305.
- and BENDER, M. LIONEL (eds.). 1981. *Nilo-Saharan: proceedings of the first Nilo-Saharan linguistics colloquium* (Leiden, 8–10 Sept. 1980). Dordrecht: Foris.
- SCHAUB, WILLI. 1985. *Babungo*. London: Croom Helm.

- SCHLERATH, BERNFRIED, and RITTER, VERONICA (eds.). 1985. *Grammatische Kategorien: Funktion und Geschichte*. Wiesbaden: Dr Ludwig Reichert.
- SCHROEDER, CHRISTOPH. 1999. *The Turkish nominal phrase in spoken discourse* (*Turcologica* 49). Wiesbaden: Harrassowitz.
- SCHWARTZ, ARTHUR. 1971. ‘General aspects of relative clause formation’. *Working Papers on Language Universals*, 6: 139–71.
- SEARLE, JOHN. 1969. *Speech acts: an essay in the philosophy of language*. Cambridge: Cambridge University Press.
- SEILER, HANSJAKOB. 1978a. ‘Determination: a functional dimension for interlanguage comparison’, in Seiler 1978b: 301–28.
- (ed.). 1978b. *Language universals: papers from the conference at Gummersbach/Cologne, Germany Oct. 3–8, 1976*. Tübingen: Narr.
- 1983. *Possession as an operational dimension of language*. Tübingen: Narr.
- 1985. ‘Kategorien als fokale Instanzen von Kontinua, gezeigt am Beispiel der nominalen Determination’, in Schlerath and Ritter 1985: 435–48.
- 1986. *Apprehension: language, object, and order. Part 3: The universal dimension of apprehension*. Tübingen: Narr.
- 1990. ‘A dimensional view of numeral systems’, in Croft et al. 1990: 187–208.
- and LEHMANN, CHRISTIAN (eds.). 1982. *Apprehension: das sprachliche Erfassen von Gegenständen. Part 1: Bereich und Ordnung der Phänomene*. Tübingen: Narr.
- and PREMPER, WALDFRIED (eds.). 1991. *Partizipation: das sprachliche Erfassen von Sachverhalten* (Language Universals Series 6). Tübingen: Gunter Narr.
- and STACHOWIAK, FRANZ JOSEF (eds.). 1982. *Apprehension: das sprachliche Erfassen von Gegenständen. Part 2: Die Techniken und ihr Zusammenhang in Einzelsprachen*. Tübingen: Narr.
- SEILER, WALTER. 1986. ‘Noun-classificatory verbal prefixes as reanalysed serial verbs’. *Lingua*, 68: 189–206.
- SELLNER, MANFRED B. 1982. ‘Strategien aus dem Japanischen für Tierklassifikatoren’. *Papiere zur Linguistik*, 26/1/82): 35–45.
- SENFT, GUNTER. 1985. ‘Klassifikationspartikel im Kilivila: Glossen zu ihrer morphologischen Rolle, ihrem Inventar und ihrer Funktion in Satz und Diskurs’. *Linguistische Berichte*, 99: 373–93.
- 1986. *Kilivila: the language of the Trobriand islanders* (Mouton Grammar Library 3). Berlin: Mouton de Gruyter.
- (p.382) SENFT, GUNTER. 2000a. ‘What do we really know about nominal classification systems?’, in Senft 2000b: 11–49.
- (ed.). 2000b. *Systems of nominal classification* (Language, Culture and Cognition 4). Cambridge: Cambridge University Press.

- SEOK CHOONG SONG. 1975. ‘Rare plural marking and ubiquitous plural marker in Korean’, in Grossman et al. 1975: 536–46.
- SERZISKO, FRITZ. 1982. ‘Numerus/Genus-Kongruenz und das Phänomen der Polarität am Beispiel einiger ostkuschitischer Sprachen’, in Seiler and Stachowiak 1982: 179–200.
- SEUREN, PIETER A. M. (ed.). 1974a. *Semantic syntax*. Oxford: Oxford University Press.
- 1974b. ‘Introduction’, in Seuren 1974a: 1–27.
- SHIBATANI, MASAYOSHI, and BYNON, THEODORA (eds.). 1995. *Approaches to language typology*. Oxford: Clarendon Press.
- SHOPEN, TIMOTHY (ed.). 1985. *Language typology and syntactic description*. Vol. i: *Clause structure*. Vol. ii: *Complex constructions*. Vol. iii: *Grammatical categories and the lexicon*. Cambridge: Cambridge University Press.
- SICHELSCHMIDT, LORENZ. 1986. ‘Optionality and choice in the analysis of adjective order: comments on Ney’. *Studia Linguistica*, 40/2: 135–48.
- SIEWIERSKA, ANNA. 1984. ‘Phrasal discontinuity in Polish’. *Australian Journal of Linguistics*, 4/1: 57–71.
- 1988. *Word order rules*. London: Croom Helm.
- 1991. *Functional Grammar* (Linguistic Theory Guides). London: Routledge.
- 1996a. ‘The distribution of subject and object agreement and word order type’. *Studies in Language*, 20/1: 115–61.
- 1996b. ‘Word order type and alignment type’. *Sprachtypologie und Universalienforschung (STUF)* 49/2: 149–76.
- 1997a. ‘Variation in major constituent order: a global and a European perspective’, in Siewierska 1997b 475–551.
- (ed.). 1997b. *Constituent order in the languages of Europe* (Empirical Approaches to Language Typology/Eurotyp 20–1). Berlin: Mouton de Gruyter.
- 1998. ‘On nominal and verbal person marking’. *Linguistic Typology*, 2/1: 1–57.
- and UHLÍŘOVÁ, LUDMILA. 1997. ‘An overview of word order in the Slavic languages’, in Siewierska 1997b: 105–49.
- RIJKHOFF, JAN, and BARKER, DIK. 1997. ‘Appendix: 12 word order variables in the languages of Europe’, in Siewierska 1997b 783–812.
- SILVERSTEIN, MICHAEL. 1976. ‘Hierarchy of features and ergativity’, in Dixon 1976: 112–72.
- SIMPSON, J. 1983. ‘Aspects of Warlpiri morphology and syntax’. Ph.D. dissertation, MIT.
- SINGLER, JOHN V. (ed.). 1990. *Pidgin and Creole tense-mood-aspect systems*. Amsterdam: John Benjamins.

- SINHA, N. K. 1975. *Mundari grammar* (Central Institute of Indian Languages Grammar Series 2). Mysore: Central Institute of Indian Languages.
- SINOR, DENIS (ed.). 1988. *Handbuch der Orientalistik, Achte Abteilung*. vol. v/1: *The Uralic languages: description, history and foreign influences*. Leiden: Brill.
- SKORIK, PETR JA. 1960–77. *Grammatika čukotskogo jazyka* (A grammer of the Chukchee language), 2 vols. Moskow: Izdatel'stvo Akademii Nauk SSSR.
- SLOBIN, DAN I. 1985. 'The child as linguistic icon-maker', in Haiman 1985b: 221–48.
- (p.383) SMITH-STARK, T. CEDRIC. 1974. 'The plurality split', in La Galy et al. 1974: 657–71.
- SOHN, HO-MIN. 1994. *Korean*. London: Routledge.
- SONG, JAE JUNG. 1991. 'On Tomlin, and Manning and Parker on basic word order'. *Language Sciences*, 13/1: 89–97.
- 1997. 'The so-called plural copy in Korean as a marker of distribution and focus'. *Journal of Pragmatics*, 27: 203–24.
- SPEISER, EPHRAIM A. 1941. *Introduction to Hurrian*. New Haven: American Schools of Oriental Research.
- SPIRO, MELFORD E. (ed.). 1965. *Context and meaning in cultural anthropology*. Glencoe, Ill.: Free Press.
- SPRUIT, ARIE. 1986. 'Abkhaz studies'. Ph.D. dissertation, University of Leiden.
- SPULER, B. (ed.). 1964. *Handbuch der Orientalistik*. vol. v/2: *Mongolistik*. Leiden: Brill.
- SRIVASTAV, V. 1991. 'The syntax and semantics of correlatives'. *Natural Language and Linguistic Theory*, 9: 637–86.
- STAMPE, DAVID. 1976. 'Cardinal number systems'. *Papers from the twelfth meeting of the Chicago Linguistic Society*: 594–609.
- STASSEN, LEON. 1985. *Comparison and universal grammar*. Oxford: Blackwell.
- 1986. 'The comparative compared'. *Journal of Semantics*, 3: 143–82.
- 1997. *Intransitive predication*. Oxford: Oxford University Press.
- STEELE, SUSAN. 1975. 'On some factors that affect and effect word order', in Li 1975: 197–268.
- STEINBERG, DANNY D., and JAKOBOVITS, LEON A. (eds.). 1971. *Semantics: an interdisciplinary reader in philosophy, linguistics and psychology*. Cambridge: Cambridge University Press.
- STENGAARD, BIRTE. 1991. *Vida y muerte de un campo semántico: estudio de la evolución semántica de los verbos latinos stare, sedere, e iacere del latín al romance del s. XIII* [Life and death of a semantic field. Study of the semantic evolution of the Latin verbs *stare*, *sedere*, and *iacere* from Latin to 13th-century Romance]. Tübingen: Niemeyer.

- STOLTENBERG, HANS L. 1950. *Etruskische Sprachlehre mit vollständigem Wörterbuch*. Leverkussen: Gottschalk.
- STROOMER, HARRY. 1987. *A comparative study of three southern Oromo dialects in Kenya: phonology, morphology and vocabulary* (Cushitic Language Studies 6). Hamburg: Buske.
- 1995. *A grammar of Boraana Oromo (Kenya): phonology, morphology and vocabulary* (Cushitic Language Studies 11). Cologne: Rüdiger Köppe.
- SUÁREZ, JORGE A. 1983. *The Mesoamerican Indian languages*. Cambridge: Cambridge University Press.
- SUN, CHAO-FEN. 1988. ‘The discourse function of numeral classifiers in Mandarin Chinese’. *Journal of Chinese Linguistics*, 16/2: 298–322.
- SWANTON, JOHN R. 1911. ‘Tlingit’, in Boas 1911b: 159–204.
- SYLLA, YÈRO. 1982. *Grammaire moderne du pulaar*. Dakar: Les Nouvelles Éditions Africaines.
- SZABOLCSI, ANNA. 1994. ‘The noun phrase’, in Kiefer and Kiss 1994: 179–274.
- SZAKOS, JÓZSEF. 1994. *Die Sprache der Cou: Untersuchungen zur Synchronic einer austronesischen Sprache auf Taiwan*. Dissertation, Rheinischen Friedrich-Wilhelms-Universität.
- (p.384) TALLERMAN, MAGGIE. 1997. ‘Word order in Celtic’, in Siewierska 1977b: 21–45.
- TAYLOR, CHARLES V. 1985. *Nkore-Kiga* (Croom Helm Descriptive Series). London: Croom Helm.
- TCHEKHOFF, CLAUDE. 1981. *Simple sentences in Tongan* (Pacific Linguistics B-81). Canberra: Australian National University.
- TEDESCHI, PHILIP J., and ZAENEN, ANNIE (eds.). 1981. *Tense and aspect* (Syntax and Semantics 14). New York: Academic Press.
- TEFERRA, A. 1987. ‘Bališša: women’s speech among the Sidama’. *Journal of Ethiopian Studies*, 20: 44–59.
- TESNIÈRE, LUCIEN. 1959. *Éléments de syntaxe structurale*. Paris: Klincksieck.
- TESTELEC, YAKOV G. 1997a. ‘Word order in Kartvelian languages’, in Siewierska 1997b: 235–56.
- 1991b. ‘Word order variation in some SOV languages of Europe’, in Siewierska 1997b: 649–79.
- 1997c. ‘Word order variation in Daghestanian languages’, in Siewierska 1997b: 257–80.
- THELIN, NILS B. (ed.). 1990. *Verbal aspect in discourse: contributions to the semantics of time and temporal perspective in Slavic and non-Slavic languages*. Amsterdam: Benjamins.

- THELWALL, ROBIN (ed.). 1978. *Aspects of language in the Sudan* (Occasional Papers in Linguistics and Language Learning 5). Coleraine: New University of Ulster.
- THOMAS, DAVID D. 1976. 'A Chrau noun phrase battery'. *Mon-Khmer Studies*, 5: 135–8. Manila: SIL.
- THOMASON, SARAH GREY, and KAUFMAN, TERENCE. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley and Los Angeles: University of California Press.
- THOMPSON, LAURENCE C. 1965. *A Vietnamese grammar*. Seattle: University of Washington Press.
- THOMPSON, SANDRA A. 1988. 'A discourse approach to the cross-linguistic category "adjective"', in Hawkins 1988b: 167–85.
- THOMSEN, M. 1984. *The Sumerian language*. Copenhagen: Akademisk Forlag.
- THORNE, BARRIE, and HENLEY, NANCY (eds.). 1975. *Language and sex: difference and dominance*. Rowley, Mass.: Newbury House.
- TIFFOU, ÉTIENNE, and PESOT, JÜRGEN. 1989. *Contes du Yasin: introduction au bourouchaski du Yasin avec grammaire et dictionnaire analytique* (Études bourouchaski 1. Asie et Monde Insulindien 16). Paris: Peeters/Selaf.
- TISCHLER, JOHANN (ed.). 1982. *Serta Indogermanica: Festschrift für Günter Neumann zum 60. Geburtstag*. (Innsbrucker Beiträge zur Sprachwissenschaft 40). Innsbruck: Institut für Sprachwissenschaft der Universität.
- TOMLIN, RUSSELL S. 1986. *Basic word order: functional principles*. London: Croom Helm.
- (ed.). 1987. *Coherence and grounding in discourse*. Amsterdam: Benjamins.
- TOMPA, JÓZSEF. 1968. *Ungarische Grammatik* (Janua Linguarum, Series Practica 96). The Hague: Mouton.
- TOPPING, DONALD M. 1973. *Chamorro reference grammar*. Honolulu: University Press at Hawaii.
- TOSCO, MAURO. 1998. 'A parsing view on inconsistent word order: articles in Tigre and its relatives'. *Linguistic Typology*, 2/3: 355–80.
- (p.385) TRAUGOTT, ELIZABETH CLOSS. 1975. 'Spatial expressions of tense and temporal sequencing'. *Semiotica*, 15: 207–30.
- and HEINE, BERND (eds.). 1991. *Approaches to grammaticalization*, 2 vols. (Typological Studies in Language 19). Amsterdam: Benjamins.
- TRUDGILL, PETER. 1992. 'Dialect typology and social structure', in Jahr 1992: 195–212.
- T'SOU, BENJAMIN K. 1976. 'The structure of nominal classifier systems', in Jenner et al. 1976: 1215–47.
- TSUNODA, TASAKU, ITOH, YOSHIAKI, and UEDA, SUMIE. 1995. 'Adpositions in word order typology'. *Linguistics*, 33/4: 741–61.

- TUCKER, ARCHIBALD N. (in collaboration with HACKETT, PETER E.). 1959. *Le Groupe linguistique Zande*. Tervuren: Koninklijk Museum van Belgisch-Congo.
- 1994. *A grammar of Kenya Luo (Dholuo)*, ed. Chet A. Creider (Nilo-Saharan 8.1). Cologne: Rüdiger Köppe.
- and BRYAN, MARGARET A. 1966. *The non-Bantu languages of north-eastern Africa (with a supplement on the Ethiopian languages by Wolf Breslau) (Handbook of African languages, part 3)*. London: Oxford University Press.
- and MPAAYEI, TOMPO OLE. 1955. *A Maasai grammar: with vocabulary*. London: Longmans, Green & Co.
- TUITE, KEVIN. 1992. ‘The category of number in Common Kartvelian’, in Aronson 1992: 245–84.
- TUNG, T'UNH-HO. 1964. *A descriptive study of the Tsou language, Formosa* (Institute of History and Philology, Special Publication 48). Taipei: Academica Sinica.
- ULTAN, RUSSELL. 1978a. ‘On the development of a definite article’, in Seiler 1978b: 249–65.
- 1978b. ‘The nature of future tenses’, in Greenberg et al. 1978: iii. 83–123.
- 1978c. ‘Toward a typology of substantival possession’, in Greenberg et al. 1978: iv. 11–50.
- UNDERHILL, ROBERT. 1976. *Turkish grammar*. Cambridge, Mass.: MIT Press.
- UNSETH, PETER. 1988. ‘Majang nominal plurals, with comparative notes’. *Studies in African Linguistics*, 19/1: 75–91.
- UNTERBECK, BARBARA. 1993. *Kollektion, Numeralklassifikation und Transnumerus*. Frankfurt am Main: Peter Lang.
- and MATTI RISSANEN. 1999. *Gender in grammar and cognition*. Vol. i *Approaches to gender*. Vol. ii *Manifestations of gender* (Trends in Linguistics: Studies and Monographs 124). Berlin: Mouton de Gruyter.
- UNTERMANN, JÜRGEN, and BROGYANYI, BELA (eds.). 1984. *Das Germanische und die Rekonstruktion der indogermanischen Grundsprache*. Amsterdam: Benjamins.
- VALDMAN, ALBERT, and HIGHFIELD, ARNOLD (eds.). 1980. *Theoretical orientations in Creole studies*. New York: Academic Press.
- VALOIS, D. 1991. ‘The internal syntax of DP’. Ph.D. dissertation, UCLA.
- VAN BAAR, TIM. 1997. *Phasal polarity* (Studies in Language and Language Use 30). Ph.D. dissertation, University of Amsterdam. Amsterdam: IFOTT.
- VAN BAARDA, M. J. 1891. *Beknopte Spraakkunst van de Galillareesche taal* [Short grammar of the Galela language]. Utrecht: Kemink & Zoon.
- (p.386) VAN BAARDA, M. J. 1908. *Leiddraad bij het bestuderen van 't Galela'sch dialekt, op het eiland Halmahera* [Manual for the study of the Galela dialect, on the island of Halmahera]. The Hague: Nijhoff.

- VAN DE KERKE, SIMON. 1996. ‘Agreement in Quechua: evidence against Distributed Morphology’, in Cremers and den Dikken 1996: 121–31.
- VAN DEN BERG, RENÉ. 1989. ‘A grammar of the Muna language’. Ph.D. dissertation, University of Leiden.
- VAN DER AUWERA, JOHAN. 1990. ‘Coming to terms’. Dissertation (*aggregatie*), Universitaire Instelling Antwerpen.
- 1998a. ‘Phasal adverbials in the languages of Europe’, in van der Auwera 1998b: 25–145.
- (ed.). 1998b. *Adverbial constructions in the languages of Europe* (Empirical Approaches to Language Typology/Eurotyp 20–3). Berlin: Mouton de Gruyter.
- and GOOSSENS, LOUIS (eds.). 1987. *Ins and outs of the predication* (Functional Grammar Series 6). Dordrecht: Foris.
- VAN DER SANDT, ROB A. 1992. ‘Presupposition projection as anaphora resolution’. *Journal of Semantics*, 9: 333–77.
- VAN LANGENDONCK, WILLY, 1995. ‘Categories of word order iconicity’, in Landsberg 1995: 79–90.
- VAN SCHAAIK, GERJAN. 1996. *Studies in Turkish grammar* (Turcologica 28). Wiesbaden: Harrassowitz.
- VAN VALIN, ROBERT D., Jr. and LAPOLLA, RANDY J. 1997. *Syntax: structure, meaning and function*. Cambridge: Cambridge University Press.
- VENDLER, ZENO. 1971. ‘Singular terms’, in Steinberg and Jakobovits 1971: 115–33.
- VENNEMANN, THEO. 1973. ‘Explanation in syntax’, in Kimball 1973b: 1–50.
- 1976. ‘Categorial grammar and the order of meaningful elements’, in Julland 1976: 615–34.
- VERKUYL, HENK J. 1972. *On the compositional nature of the aspects*. Dordrecht: Reidel.
- 1993. *A theory of aspectuality*. Cambridge: Cambridge University Press.
- VET, CO. 1986. ‘A pragmatic approach to Tense in Functional Grammar’. *Working Papers in Functional Grammar*, 16 (Universiteit van Amsterdam).
- VIDAL, ALEJANDRA, and KLEIN, HARRIET E. MANELIS. 1998. ‘Irrealis in Pilagá and Toba? Syntactic versus pragmatic coding’. *Anthropological Linguistics*, 40/2 (special issue: ‘Symposium or Irrealis’): 175–97.
- VILKUNA, MARIA. 1997. ‘Word order in European Uralic’, in Siewierska 1997b: 173–233.
- VISSE, LEONTINE E., and VOORHOEVE, C. L. 1987. *Sahu–Indonesian–English dictionary and Sahu grammar sketch* (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 126). Dordrecht: Foris.
- VOGEL, PETRA M., and COMRIE, BERNARD (eds.). 2000. *Approaches to the typology of word classes*. Berlin: Mouton de Gruyter.

- VONEN, ARNFINN MURUVIK. 1994. 'Multifunctionality and morphology in Tokelau and English'. *Nordic Journal of Linguistics*, 17/2: 155–78.
- VON GARNIER, K. 1909. 'COM- als perfektierendes Praefix bei Plautus, SAM- im Rigveda, CYN- bei Homer'. *Indogermanische Forschungen*, 25: 86–109.
- (p.387) VON HEUSINGER, KLAUS. 1999. 'Spezifität als semantische Kategorie'. Habilitationsvortrag, Universität Konstanz.
- VOORHOEVE, C. L. 1965. *The Flamingo Bay dialect of the Asmat language* (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 46). The Hague: Nijhoff.
- VŨ DUY-TÙ'. 1983. *Lehrbuch der vietnamesischen Sprache: eine Einführung mit Übungen, Lösungen und Tonmaterial*. Hamburg: Buske.
- WACKERNAGEL, JOHANN. 1892. 'Über ein Gesetz der indogermanischen Wortstellung'. *Indo-Germanische Forschungen*, 1: 333–436.
- WANG FU-SHIH. 1972. 'The classifier in the Wei Ning dialect of the Miao language in Kweichow', in Purnell 1972: 111–86.
- WATKINS, LAUREL J. 1976a. 'Position in grammar: sit, stand, lie'. *Kansas Working Papers in Linguistics*, 1: 16–41.
- 1976b. 'Shape vs. position: classificatory verbs in North America'. Paper presented at the 1976 LSA annual meeting, Philadelphia.
- WATSON, SAUNDRA K. 1976. 'The Pacöh noun phrase'. *Mon-Khmer Studies*, S: 219–31.
- WAUGH, LINDA R., and MONVILLE-BURSTON, MONIQUE (eds.). 1990. *On language: Roman Jakobson*. Cambridge, Mass.: Harvard University Press.
- WEBER, DAVID JOHN. 1989. *A grammar of Huallaga (Huánuco) Quechua*. Berkeley and Los Angeles: University of California Press.
- WEIR, E. M. HELEN. 1990. 'Incorporation in Nadëb', in Payne 1990: 321–63.
- WELMERS, WILLIAM E. 1973. *African language structures*. Berkeley and Los Angeles: University of California Press.
- WERNER, HEINRICH. 1994. *Das Klassensystem in den Jenissej-Sprachen* (Veröffentlichungen der Societas Uralo-Altaica 40). Wiesbaden: Harrassowitz.
- 1995. *Zur Typologie der Jenissej-Sprachen*. Wiesbaden: Harrassowitz.
- 1997a. *Die ketische Sprache* (Tunguso-Sibirica 3). Wiesbaden: Harrassowitz.
- 1997b. *Abriß der kottischen Grammatik* (Tunguso-Sibirica 4). Wiesbaden: Harrassowitz.
- 1997c. *Das Jugische (Sym-Ketische)* (Veröffentlichungen der Societas Uralo-Altaica 50). Wiesbaden: Harrassowitz.
- WERNER, O. 1984. 'Morphologische Entwicklungen in den germanischen Sprachen', in Untermaier and Brogyanyi 1984: 181–226.

- WESTERMANN, DIEDRICH. 1912. *The Shilluk people: their language and folklore*. Westport, Conn.: Negro Universities Press. [Repr. 1970.]
- 1930. *A study of the Ewe language*. London: Oxford University Press [translation of *Grammatik der Ewe-Sprache*, 1907]. [Repr. 1952.]
- and BRYAN, MARGARET A. 1970. *The languages of West Africa (Handbook of African languages, part 2)*. London: Dawsons.
- WETZER, HARRIE. 1992. “Nouny” and “verby” adjectivals: a typology of predicative adjectival constructions’, in Kefer and van der Auwera 1992: 223–62.
- 1996. *The typology of adjectival predication* (Empirical Approaches to Language Typology 17). Berlin: Mouton de Gruyter.
- WHALEY, LINDSAY J. 1997. *Introduction to typology: the unity and diversity of language*. Thousand Oaks, Calif.: Sage.
- WHEATLEY, JULIAN K. 1987. ‘Burmese’, in Comrie 1987: 834–54.
- (p.388) WHINNOM, KEITH. 1980. ‘Creolization in linguistic change’, in Valdman and Highfield 1980: 203–12.
- WHORF, BENJAMIN LEE. 1946. ‘The Hopi language, Toreva dialect’, in Osgood 1946: 158–83.
- WIERzbicka, ANNA. 1985. “Oats” and “wheat”: the fallacy of arbitrariness’, in Haiman 1985b: 311–42.
- WILKINS, DAVID P. 1989. ‘Mparntwe Arrernte (Aranda): studies in the structure and semantics of grammar’. Ph.D. dissertation, Australian National University, Canberra.
- 2000. ‘Ants, ancestors and medicine: a semantic and pragmatic account of classifier constructions in Arrernte (Central Australia)’, in Senft 2000b: 147–216.
- WILLETT, THOMAS. 1988. ‘A cross-linguistic survey of the grammaticalization of evidentiality’. *Studies in Language*, 12/1: 51–97.
- WILLIM, EWA. 1995. ‘The syntax of noun phrases in Polish: linearization parameters’, in Palek 1995: 166–89.
- WILTSHERE, C., and GRACZYK, R., and Music, B. (eds.). 1989. *CLS 25: papers from the 25th annual regional meeting of the Chicago Linguistic Society*. Part 1: *The general session*. Chicago: Chicago Linguistic Society.
- WOODBURY, HANNI JOCH. 1975a. ‘Onondaga noun incorporation: some notes on the interdependence of syntax and semantics’. *International Journal of American Linguistics*, 41/1: 10–20.
- 1975b. ‘Noun incorporation in Onondaga’. Ph.D. dissertation, Yale University.
- WURM, STEPHAN A. 1982. *Papuan languages of Oceania* (Acta Linguistica 7). Tübingen: Narr.
- XRAKOVSKIJ, VIKTOR S. (ed.). 1989. *Tipologija iterativnyx konstrukcij* [Typology of iterative constructions]. Leningrad: Nauka.

- YNGVE, VICTOR H. 1960. 'A model and an hypothesis for language structure'. *Proceedings of the American Philosophical Society*, 104: 444–66.
- YOO, HYUN-KYUNG. 1997. 'A study of Korean adjectives'. Ph.D. dissertation, Yonsei University, Seoul.
- ZABORSKI, ANDRZEJ. 1976. 'Cushitic overview', in Bender 1976: 67–84.
- 1986. *The morphology of nominal plural in the Cushitic languages* (Veröffentlichungen der Institute für Afrikanistik und Ägyptologie der Universität Wien 39. Beiträge zur Afrikanistik 28). Vienna: Institut für Afrikanistik.
- ZEPEDA, OFELIA. 1983. *A Papago grammar*. Tucson: University of Arizona Press.
- Z'GRAGGEN, J. A. 1965. 'Possessor–possessed relationships in the Saker language, northeastern New Guinea'. *Oceanic Linguistics Monographs*, 4: 119–26.
- ZIMMERMANN, ILSE (ed.). 1991. *Syntax und Semantik der Substantivgruppe* (Studia Grammatica 23). Berlin: Zentralinstitut für Sprachwissenschaft.
- ZIPF, GEORGE KINGSLEY. 1935. *The psycho-biology of language: an introduction to dynamic philology*. Boston: Houghton & Mifflin.
- ZIV, YAEL, and COLE, PETER. 1974. 'Relative extraposition and the scope of definite descriptions in Hebrew and English'. *Papers from the tenth regional meeting of the Chicago Linguistic Society*: 772–86.
- ZUBIN, DAVID A., and KÖPCKE, KLAUS-MICHAEL. 1986. 'Gender and folk taxonomy: the indexical relation between grammatical and lexical categorization', in Craig 1986b: 139–80.

(p.389) Index of Subjects

- abstract noun 19, 100, 171, 278, 341
actual/non-actual distinction 136, 232–8
adjectival notion 132–4, 136, 138–9, 145, 274, 289, 299, 305
adjective 1–2, 4, 8, 10–12, 14–15, 17–18, 24, 57, 73, 77–8, 80, 85–6, 93–4, 96–100, 111, 121–45, 152, 168, 171, 173, 175, 177, 179, 182–3, 192, 196–7, 204, 212, 218–20, 223–5, 228, 254, 257, 261–79, 284–5, 287–90, 298–305, 308–11, 313–23, 327–32, 334, 336–7, 341, 343, 345–9
adjective-noun adjacency 266–76, 322
adjective-noun order 129, 133
adjective-relative clause order 298–9
attributive adjective appearing with copula 125–6
attribute/predicate distinction 125, 228, 269
'big'-fronting 267
doubling 140–1
Greenberg's Universal 18 304–5
Greenberg's Universal 20 273–6
incorporation 132–3, 135, 138
lgs. with adjectives 122–9, 266–76
lgs. with minor or closed class of adjectives 18, 129–33, 143, 270–3, 301, 304, 323
lgs. without adjectives 133–9, 284, 289
necessary condition 141–5
stacking 218
adjective phrase 264, 310
adposition, *see* relator
adpositional phrase 209–11
adverb 15–16, 22, 31, 33, 80, 104, 111, 157, 167, 171, 214–17, 223–4, 230, 233, 256, 264, 279, 285, 287–8, 298, 329
adverbial 16–17, 93, 111, 165, 175, 177, 206, 211, 214, 217, 236, 255, 258, 279, 281, 298
affix 12, 22, 37, 46, 51, 63, 68, 70, 73–4, 79–80, 83, 93, 98–9, 102–3, 106, 111, 118–21, 123, 127, 130, 132, 148, 150, 166, 180–1, 192, 197, 233, 240, 244–5, 270, 279, 306, 318, 340
infix 138
prefix 12, 14, 34–5, 44, 69, 71, 81, 85, 88–92, 94, 96, 98, 108–9, 112, 115, 120, 128, 131–2, 138, 151, 158–9, 183, 189, 196–7, 240–1, 301, 305–6
suffix 12, 18, 20, 32–7, 43–5, 51, 62, 71–3, 79–80, 88, 90–1, 93–4, 98, 102, 106, 108–10, 112–15, 118–20, 124, 127–8, 130, 132, 136, 138, 140, 143–4, 149, 153–4, 158–9, 161, 164, 166–8, 170, 181–2, 186, 189–90, 192–3, 196–9, 202–3, 209–11, 233, 236, 240, 241, 243, 246, 257, 268–9, 271, 274–6, 301, 304–6, 317–18, 322–3, 326, 332
agentive suffix 119, 235
agglomerative 52–4, 147
 see also homogeneity
agreement 20, 26, 34–8, 60–1, 81–2, 85, 96–8, 105–6, 108–9, 112–13, 123–8, 131–2, 138, 144, 163, 167, 171, 182, 240–2, 244, 246–7, 257, 270, 275, 298, 311, 318, 339, 347
Aktionsart 58–9, 101, 104, 338, 340, 346
alignment system 247
 ergative–absolutive 71
 nominative–accusative 71

split ergativity 71
ambiguity 51, 102, 105, 147, 165, 206, 209, 235, 246, 310, 334–5
anchoring, *see reference*
animate/inanimate distinction 30–1, 35–6, 41, 43, 51, 60–99 (esp. 86–91), 112, 114–15, 152, 181, 193, 196, 202
apposition 1, 19–23, 32, 39, 73, 99, 126, 133, 139–40, 160, 162–5, 168–9, 180, 183, 189, 191–2, 194, 200, 204–7, 245, 255–7, 270–6, 278, 283–4, 298, 319, 326–7, 330, 340, 342
article 1, 3, 15, 22–3, 43–4, 57, 62, 69, 80, 84, 92–6, 98–9, 109, 119–21, 150, 155, 163, 166, 173–4, 176, 178–9, 182–94, 204, 229, 235, 238, 254, 261, 267–8, 272, 274–5, 279–80, 285–6, 298, 319–22, 336–7, 342
(p.390) aspect 4, 12, 27, 31, 38, 42–3, 48, 57–9, 65, 74, 79, 86, 100, 101–21, 130, 142, 145–7, 149–53, 155–6, 166, 170, 173, 178, 185, 196, 200, 212, 214–21, 224–5, 228, 239, 244, 253, 268, 313, 315–18, 320–3, 337, 339–41, 346–7
 see also nominal aspect; verbal aspect
aspect marking 31, 38, 42–3, 59, 65, 79, 86, 100–1, 103–6, 108, 112–13, 115, 117–21, 142, 145, 147, 149, 151–3, 155–6, 166, 170, 173, 212, 215, 218–20, 221, 268, 313, 315–16, 318, 320–3, 337, 340–1, 346–7
asymmetry
 see also symmetry
pre-/postnominal modification 21, 165, 273–6, 311, 318
pre-/postnominal relative clause 295, 311, 346
pre-/postverbal object clause 294–5
NP/clause structure 234
augmentative 84, 152, 329
auxiliary verb 282–3
backgrounding 58, 177
 see also topicality
Beginning (feature) 57–8, 103–5, 121, 166, 198, 215–16, 219, 266, 296, 308, 322, 343
body part 67, 81–2, 86, 113, 157, 197, 199, 245, 271, 306
body part counting system 34
borrowing 12, 43, 91, 108, 125, 131–2, 151, 171–2, 208, 328
Branching Direction Theory (Dryer) 266, 278–91, 297–8
case 21, 204, 210
 ablative 70–1, 96, 124, 209
 absolutive 33, 38–9, 109, 132, 160, 181, 189, 191, 330
 accusative 42, 71, 96, 112
 allative 70–1
 autonomous 112, 139
 dative 71, 96
 ergative 26, 71, 189, 191
 essive 43
 genitive 21, 73, 79, 96, 131, 166, 194–205, 256, 259, 261, 277, 279, 298, 310, 334, 345
 instrumental 3, 43, 109
 locative 70–1, 79, 86, 90, 93, 109, 112
 nominative 81, 96, 112, 164
 oblique 132
 vocative 165
case marker, *see* relator

classification
languages 5–8
nominal subcategories 54, 338; *see also Seinsart*
parts-of-speech 16–17, 287–9
classifier 2, 4, 21, 26, 29, 34, 38–9, 41–2, 45, 47–53, 55, 60, 63, 68–9, 71, 73–8, 84, 90–1, 99, 102, 115–17, 119, 121, 135, 141–4, 147, 153, 157–66, 179, 184–5, 191, 254–5, 270, 284–5, 331–3, 336–40, 342
see also noun classification
anaphoric function 77, 165, 191, 270
classifier–numeral order 39, 163, 255, 260
clf + num–noun order 157–9, 254–5, 260
common classifier 48
collective classifier 48, 116–17, 119, 142
collocating constituents 78, 165–6
diachronic development 141–5, 165, 186, 341
discourse function 191
doubling of num + clf 164–6
general classifier 49–50, 77, 141, 333, 338
genitive classifier 73
individualizing function 50
intra-locative classifier 71
mensural classifier 47, 49, 52
multiple class membership 69
noun classifier 39, 73–5, 78, 99, 285
number of classifiers 77, 165
numeral classifier, *see* sortal classifier
optionality of classifier 77–8
other functions 38, 165
possessive classifier 21, 73, 90–1
properties of prenominal classifiers 165–6
relational classifier 90
sortal classifier 2, 21, 29, 39, 41, 47–52, 73–8, 84, 99, 102, 115–17, 119, 121, 141–5, 153–5, 157–9, 161–6, 184–6, 191, 254–5, 260, 284, 331, 333, 340
variation in classifier choice 76, 91
classificatory article 69
classificatory element in verb 69–70
classificatory particle 270
classificatory verb 68–70, 76
clitic 31, 32, 73–4, 120–1, 130, 180, 191–2, 202, 204, 209, 245, 249, 258, 286–7, 294, 317, 319, 321
clause structure 214–38
 discourse modifier 229, 231–8, 336, 347
(p.391)
dual function of predication 231–8
layered structure 216–18, 224, 237, 344, 348
localizing modifier 214–18, 224, 237
qualifying modifier 214–18, 224, 237
quantifying modifier 214–18, 224, 237
code–quantity principle (Givón) 25
cognitive factors 3, 55, 178, 223–5, 343

Cognitive Grammar 228
collective aspect marking, *see* nominal aspect
collective noun, *see* noun
collectivity 37, 42–4, 48, 51, 53–5, 72, 79–80, 86, 100–21, 141–2, 144–7, 149–50, 153–4, 170, 181, 219–20, 286, 316–19, 322–3, 329, 338–42, 347
colloquial language, *see* spoken/written language
communication 25, 28, 36, 38, 43, 80, 108, 111–12, 126, 132, 134, 139, 160, 204, 228, 299, 301, 310, 322, 324, 330–1
communicative function 3, 11
comparative construction 289–90
competing motivations 259
complement 51, 232, 244, 266, 281–2
complementizer, *see* relator
complexity 20, 23–7, 199, 208–9, 223–4, 245, 249, 252, 257, 275, 294–5, 307, 335, 397
 coordination of modifiers 25
culture 25
distribution of modifiers over NPs 24–5
distribution of NPs over clauses 25, 312
men’s/women’s language 26
prenominal/postnominal position 20
style 24, 26
taboo situations 23
compulsory 32, 34–5, 43, 45, 50, 73–4, 77, 101–2, 104, 106, 108, 110–11, 114, 119, 121, 132, 147, 153–4, 172, 186, 192, 200, 312, 340, 348
concord, *see* agreement
configurational/non-configurational 20
conjunction, *see* relator
content variable, *see* predicate variable
contiguity 266–7, 309, 311, 322, 348
co-occurrence restriction 329–31
coordinating conjunction, *see* relator
coordination 25, 245, 249, 251, 290, 294–6
copula 61, 64–5, 69, 125, 134–5, 168–70, 269, 279, 282, 298, 345
correlation 25–6, 49, 56, 60–1, 68, 70–2, 79, 81, 85, 87, 94, 97, 104, 144, 176, 189, 198, 230, 240, 278–86, 297–8, 302, 304–7, 336, 345
counter example (real or apparent) 144, 153, 255–9, 266–76, 276–7, 299, 304–7, 317, 319, 321, 325, 327, 345–6
coverb 245–6
cross-referencing 22, 114, 194–6, 200–2, 229, 240–2, 246–7, 302
cultural factor 23–6, 61, 79, 110
cumulative 52, 147
 see also homogeneity
definite article 69, 80, 93–4, 120, 166, 174, 176, 178–9, 182, 185–9, 191, 238, 272, 275, 286, 337
definiteness 36–7, 78, 94, 108, 113, 116, 130, 140, 153, 165, 177, 182, 185–6, 188–92, 198, 254, 275, 286, 305, 326, 340
definiteness marking
 classifier 191
 demonstrative pronoun 188–9

distinct form 189–91
possessive affix 186
stress 185
word order 185

deixis 25, 93, 95, 120, 136, 166, 173–6, 178, 185, 188, 214, 217, 229, 231, 257
discourse 231
social 72
spatial 173, 178–85
temporal 214
weak 173, 229

demonstrative pronoun 2, 22, 73, 76–9, 85, 92–4, 97, 99, 116, 120, 141, 165–6, 173–6, 178–86, 188–9, 191–2, 199, 208, 219, 224, 247–8, 254, 267–74, 275–6, 278, 284–5, 290–1, 299–305, 308, 311, 313–16, 323–30, 332–4, 336–7, 340, 342, 348, 435
bound form 180–1
co-occurring with article 182–4, 254
co-occurring with classifier 184–5
doubling 181–2
Greenberg's *Universal* 18 304–5
Greenberg's *Universal* 20 273–6

dependency reversal 168

descriptive/referential function of NP 228–38

diachronic development 12, 23, 31, 37, 44, 48, 63, 65, 68–71, 73, 76, 78, 80, 83–5, 92–5, 97, 102, 104, 108, 116–17, 119–21, 124, 131–2, 141, 149, 151–2, 154, 156, 164–7, 174, 178, 180, 182, 186, 188, 190–1, 193–4, 198, 224, 240, 254–5, 326, 340–1

(p.392) dimension

space 19, 28, 43–4, 48, 50–7, 59, 61–3, 71, 79, 81, 87, 100–1, 103, 105, 118, 121, 141, 145–7, 162, 167, 173–6, 178, 185, 214, 218–20, 223–4, 227–8, 231, 234, 247, 253, 276, 319, 336–41, 343–4, 347
time 19, 56–9, 61, 101, 103–5, 174–5, 177, 185, 214–19, 223–4, 228, 232–4, 340, 344

diminutive 37, 84, 118–19, 317, 323

discourse modifier NP 226–7, 229–38, 336, 347
discourse operator 173, 226–38, 342;
see also (in)definiteness, generic and (non-)specific reference
discourse satellite 226–38

discourse world 228–9, 233–5

dissectiveness 51–2
see also homogeneity

distribution 81, 94, 111–12, 123, 154, 190, 217, 229, 317, 323
categories over word classes (Dryer) 285
features over nominal subcategories 52–4
languages over basic word orders 281
languages over families 5–8
modifiers over NPs 24–5
nominal subcategories 29–31
NPs over clauses 25, 312
parts-of-speech systems 17
pattern 291
predicates over parts-of-speech 17

diversity value (DV) 5, 7–8

dummy element 23, 165, 249
dynamic verb 103
Early Immediate Constituents (Hawkins) 291–9, 345
elicitation 24, 181
embedded structure 2, 24, 33, 134, 169, 172–3, 176–7, 207, 222, 225, 250, 253–4, 258, 261, 263–5, 270, 274–6, 278, 283–5, 290–1, 293, 301, 305–11, 314, 322, 324–5, 327, 332–5, 341, 344–6, 348–9
emphasis 51, 58–9, 73, 76, 78–9, 89, 94, 110–11, 118, 126, 155, 161–2, 167, 181–2, 185, 189, 192, 194–9, 205, 236, 246, 250, 258–9, 267–8, 272–3, 285, 301, 319, 322–3, 329, 335–6, 338, 341
Ending (feature) 36, 43, 57–8, 80–1, 83, 97–8, 103–5, 111, 121, 165, 171, 215–16
entity
 discourse 173, 176, 227, 229–30, 232, 337
 first order 19, 28–9, 60–1, 102, 143–4, 212, 220, 230, 234, 287, 291, 336
 second order 19, 61, 234
 third order 19
event dominated language 26
existential construction 63–5, 86, 174, 192
extraposition 205, 259, 296
expression rule 226
exception 38, 45, 70, 91, 96, 114, 116, 140, 144, 248, 250, 262, 271, 277, 289, 304–5, 321
familiarity 23, 25–6, 151, 176–7, 223
 see also topicality
finite/non-finite distinction 208, 242, 310
focality 20, 76, 94, 120, 154–5, 223, 250, 295, 323
foregrounding 131, 209
 see also focality
Functional Grammar 2–5, 13–14, 103, 212–38, 247–8, 253, 260, 280–98, 313, 337, 344
gender 21, 31, 34–6, 42–3, 60–1, 70–1, 73–4, 78–86, 92–4, 96–7, 99, 102, 113, 116, 125, 150, 163, 171, 182, 186, 188, 193–4, 239, 257, 275, 328, 339–40
 see also noun class
 acquisition 73, 93–4, 340
 affectionate significance 84
 biological gender 61
 common gender 71, 80, 84
 diminutive significance 84
 feminine gender 21, 32, 35, 61, 70–1, 80–1, 83–6, 96–7, 99, 102, 120, 166, 193
 free gender 84
 gender replacement 82–4
 gender switch 84
 grammatical gender 43, 60–1
 indefinite gender 83–4, 193–4
 masculine gender 32, 35, 61, 70–1, 80–1, 83–6, 96–7, 99, 102, 118, 120, 128, 166, 193
 natural gender 60
 neuter gender 35, 61, 70–1, 80–1, 84–6, 91–2, 96, 99, 107, 111, 124, 138, 193
 pejorative significance 84
 phonetic gender 85
 plural gender 85
 polarity 85–6

semantic gender 60
sex gender 61

(p.393) Generative Grammar 4, 280, 293–6, 299

general noun, *see* noun

generic noun 40, 75, 111, 168

generic reference, *see* reference

gesture 33, 156

grammatical noun modifier, *see* operator

grammaticalization 23, 31, 73, 78, 116, 141, 143, 152, 254, 340
see also diachronic development

headedness 168

heaviness 205, 218, 220, 249, 252, 259, 275, 284, 292, 307, 309, 311, 329, 335, 346, 349

hierarchy

- alienable/inalienable 86
- clause structure 4, 214–18, 223–4, 231–8, 344
- domain 264–5, 282, 296, 337, 342, 345
- humanness/animacy 112–13
- NP structure 2, 19–20, 218–38, 313, 337, 342–3, 347
- number 112

Postpositional Noun Modifier Hierarchy (Hawkins) 277

Prepositional Noun Modifier Hierarchy (Hawkins) 277

proximity 253

relativization 176

historic development, *see* diachronic development

Homogeneity (feature) 28, 44, 50–7, 59, 62, 78, 104–5, 117, 121, 144, 146–7, 219, 228, 338–9, 347

honorific nouns and verbs 72

human/non-human distinction 30–1, 34, 36–9, 55, 60–99 (esp. 79–80), 106–7, 110, 112–15, 118, 123, 127–8, 140, 144, 151, 154, 159, 161, 163–4, 169, 178, 196, 201–2, 223, 228, 253, 262, 291–2, 294–5, 308, 317–18, 343, 345

iconicity 74, 216, 240, 248, 253–60, 265, 281, 295, 307, 344

ideophone 12, 91

illocution 214, 231–2, 234

- exclamation 267, 275
- imperative 233–4
- interrogative 136

illocutionary modifier 226

illocutionary verb 233

impossible 3, 56, 61, 139, 244–5, 311, 348

incorporation 28, 38, 56–7, 66–70, 91–3, 95, 122, 132–3, 155, 185, 195, 208, 225, 244, 246, 330, 339, 347

indefinite article 3, 43, 57, 80, 84, 94–5, 109, 173, 186–8, 192–4, 235, 238, 267–8, 319, 321–2

indefiniteness 3, 20, 32, 36, 39, 43, 57, 80, 83–4, 93–5, 108–9, 113, 116, 124, 127, 161, 165, 170, 173, 182–6, 188–9, 192–4, 228–30, 234–5, 238, 267–9, 319–22, 337, 341

‘iron horse’ phenomenon 25

kinship 13, 67, 71, 86, 88, 92–3, 114–15, 120, 197, 199, 202, 223, 271, 306, 318

language change, *see* borrowing; diachronic development

language processing 24, 55, 254, 258, 265–6, 291, 294–5, 308, 311, 348

language sample 5–10, 291
layered structure NP, *see* noun phrase structure
layered structure clause, *see* clause structure
lexicalization 12, 113, 135, 146, 149–51, 323
lexical noun modifier, *see* satellite
likepartedness 51–2
 see also homogeneity
linguistic diversity 5
linker, *see* relator
loan, *see* borrowing
localist hypothesis 175, 178, 223
 localizing modifier (NP) 173–212, 216–29, 233, 239, 337, 341–2, 346–7, 349
 localizing operator 173, 176, 178–94, 212, 219, 221, 225, 228, 238, 313, 333–4, 342;
 see also demonstrative pronoun
 localizing satellite 173, 175–8, 194–212, 217, 219, 221–2, 231, 233, 313, 332–4, 342–3;
 see also relative clause; possession: possessor NP/pronoun
location layer 227, 237, 313, 337, 343–4, 347
locative construction 63–4, 86, 174, 192
mass noun, *see* noun
maxims of conversation 24–5
men’s/women’s language 26
metaphor 174, 178, 219, 223–4, 343
minor noun class 84
minor noun phrase 21
minor word class 129, 286
 adjective 18, 129–33, 143, 270–3, 301, 304, 323
 noun 12–13, 92
(p.394) modal verb 233
modi essendi 55
modifier
 see also operator; satellite
bound/free distinction 31, 33–4, 38–9, 132–3, 135, 138, 159–60, 169, 180–1, 197–9
descriptive/discourse distinction 213, 228–38, 336, 347
embedded/non-embedded distinction 283–5, 290–1, 308–11
form/function distinction 223, 332, 343
grammatical/lexical distinction 278, 283–5, 291
prenominal/postnominal distinction 276–99
mood/modality 12, 169, 214–15, 232–8, 251
morphology 13, 18, 23, 26–7, 31, 33, 50, 58–60, 73–4, 86–7, 91–
4, 96, 99, 104, 110, 121, 123, 131–2, 134, 152, 171–2, 186, 194, 215–
16, 220, 226, 240, 245, 248, 253, 262, 274, 282, 288, 311, 316, 339, 346, 348
 compounding 13, 37, 49, 133, 135, 151, 160, 169, 242, 273, 316, 322, 325
 declension 96–7, 164
 derivation 14, 18–19, 37, 54, 56, 65, 69, 84, 86–7, 91, 121, 130–
2, 135, 138, 151, 163–4, 166, 170–1, 174, 192, 224, 266–7, 269–
70, 289, 323, 328, 336
 inflection 10, 12, 22, 57, 84, 121, 125–6, 128, 162, 198, 248, 323
 isolating language 193, 248
 nominalization 19, 20, 67, 73, 91, 111, 134, 135, 138, 139, 165, 186, 205, 211, 269
 polysynthetic language 38, 40, 123, 138, 164, 218, 245, 248, 311, 323, 329, 348

portmanteau element 36, 39, 71, 108, 186
reduplication 17, 37, 111, 114–15, 129, 132, 137, 148–9, 152–3
suppletion 37, 64, 112, 128, 151, 166
negative 232, 234
auxiliary 279, 283, 298, 345
correlation 25
object 93
quality/property word 135
nominal aspect 31, 38, 42, 59, 100–21, 142, 147, 149, 151–2, 155–6, 166, 173, 185, 212, 218–29, 268, 313–23, 337, 340, 346–7
collective aspect 42–4, 48, 79–80, 86, 100–21, 142, 145, 150, 153, 170, 317–18, 322–3, 340
morpho-syntax of nominal aspect marker 120–1, 316–23
origin of nominal aspect marker 119–20
singulative aspect 43, 101–4, 106, 108–9, 118–21, 148–50, 317, 319–21, 323, 340
size and nominal aspect 117–19
nominal subcategories 28–59, 62, 79, 141, 146–7, 149, 211, 239, 338–9
see also noun; *Seinsart*
definitions 53–6
incorporated noun 56–7
predicate noun 56
nominoid 91
noun (*Seinsart*)
collective noun 19, 51–5, 101, 104–5, 116, 149, 338–42
general noun 28, 49–5, 52, 54, 59, 116, 141, 145–7, 338–9, 347
mass noun 48–9, 51–5, 59, 62, 76, 141, 146, 338–9
set noun 28, 29, 30–2, 38–41, 46–7, 50, 52–4, 59, 79, 100–19, 141–5, 147–53, 155, 339–42
singular object noun 28–9, 31, 34–8, 44–5, 50–5, 59, 62, 79, 101–2, 104, 108, 116, 143, 145, 146–52, 155, 338–9, 341, 347
sort noun 28–9, 30–1, 38, 41–2, 44, 48–55, 59, 116–17, 141, 143–8, 153–5, 162–6, 338–42, 347
noun class 21, 26, 30, 34–6, 42, 60–90, 91, 94, 96–7, 99, 102, 106–17, 124, 128, 130, 150, 163, 167, 171, 190, 268, 286, 340
see also gender
alienable/inalienable possession 60, 62, 72, 86–91, 93–4, 99, 151, 198, 271, 325
declension class 96
exhaustive noun class system 80–6
humanness and animacy 79–80
morphological noun class systems 91–6
noun marker 73, 92–7, 119, 184
phonological noun class system 96–9
tense class 190
noun classification 60, 63, 67, 74–8, 82, 86, 96, 211, 270, 339
see also classifier
noun phrase function, *see* descriptive/referential function; pragmatic function; reference; semantic function; syntactic function
noun phrase marker 39, 95, 120, 183–4, 285
noun phrase structure 19–27, 213–38
see also apposition

constituency 23, 253, 293–4, 344
dependency 245, 344

(p.395)

discontinuity 20, 23, 161, 205, 239–40, 245, 250, 253, 255–60, 275, 293–4, 296
discourse modifier, *see above*
domain integrity 239–60
flat/hierarchical distinction 20
fractured NP 20, 255
fuzzy NP 256
head proximity 261–312
hierarchical structure 2, 19–20, 218–38, 313, 337, 342–3, 347
integral/non-integral NP distinction 1–2, 8, 19–
23, 27, 71, 79, 99, 110, 126, 133, 140, 157, 164–5, 167, 169, 194, 205–
6, 213, 238, 247, 254–8, 260, 272, 275–6, 278, 283–6, 291, 295, 301, 314, 316–
17, 324, 326–8, 331–2, 336, 340, 342, 345–6, 348
layered structure 2, 4–5, 73, 218–38, 313–15, 336–7, 342–4, 347–8;
 discourse modifier, *see above*;
 localizing modifier, *see above*;
 qualifying modifier, *see below*;
 quantifying modifier, *see below*
miniature NP 21
NP-operator 227
NP-satellite 227
NP-variable 227–38
nucleus 216, 227, 238, 343–4, 347
parataxis 329
referential modifier, *see* discourse modifier
scope 313–35
simple/complex distinction 23–7, 213–38, 260, 276, 278, 284, 290–1;
simple NP 2, 5, 171, 196, 205, 225, 254, 258, 260, 265, 314, 324, 328, 332, 345–6;
 complex NP 2, 24, 168, 222, 225, 260, 301, 309, 314, 335, 346, 348–9
tight/scrambled distinction 20, 23, 139, 204
NP-PP order 297
null, *see* zero
number 10, 19–22, 29, 31–2, 34–46, 50, 53, 56–7, 78–9, 81–7, 94, 101–
21, 123, 125, 127–8, 130, 146–56, 166, 170, 172, 186, 189, 191, 194, 196, 219–21, 224–
5, 246, 257, 275, 284, 326, 337–8, 340–1, 347–8
 discord 43, 56, 103, 105–14, 155, 340, 347
dual 111, 114, 146, 152–3, 170, 198
evolution 152
nominal aspect marking 155–6
plurality 29–46, 48, 50–1, 53–4, 64, 68, 79, 81–2, 84–6, 94, 97–103, 105–6, 108–
16, 118–20, 127–8, 130, 146–7, 149–55, 166, 169–71, 183, 189–90, 193–
4, 198, 225, 279, 285–7, 298, 316–17, 319, 322–3, 326, 337, 339, 341
sort nouns 153–5
verbs and nouns 44, 152
numeral 2, 21–2, 31–9, 41–3, 45–7, 50–1, 53, 57, 70, 73–80, 84–5, 97, 99–102, 105–
6, 108–15, 117, 119, 121, 134, 136, 141–3, 146–7, 149–50, 153–73, 181, 184–6, 188, 190–
2, 199, 204, 219, 221, 224, 236, 247–8, 254–6, 258, 263, 267–71, 273–4, 276, 278, 283–
7, 290–1, 300–5, 308, 311, 313–16, 318–19, 323–32, 336–42, 345–8

bound form 31, 33–4, 38–9, 159–60, 169
cardinal numeral 2, 35, 53, 75, 77, 100, 110, 146–7, 156–
68, 170, 181, 225, 242, 267, 339, 342
complex numeral 36
discontinuity 258
doubling 160–2
gesture 33, 156–7
Greenberg's *Universal* 18 304–5
Greenberg's *Universal* 20 273–6
indefinite article 186–8
numeral adverb 22, 33, 80, 111, 171
numeral classifier, *see* classifier
numeral noun 33, 165, 169–72, 326
numeral–noun order 157–9
numeral verb 36, 168–9, 341
ordinal numeral 78, 157–9, 166–8, 337
predicative numeral 33, 39, 70
Object 13, 19, 25–6, 38, 71, 90, 93, 98, 112, 241–5, 249, 252, 257–8, 276, 278–83, 294–
5, 297–8, 302–4, 330, 338, 345–7
object-dominated languages 26
Object patterner (Dryer) 279–91
obligatory, *see* compulsory
ontology 10, 27–8, 44, 55, 60, 62, 144, 230, 336, 339
 ontological classification of entities 55
 ontological correlate/counterpart 10, 27–8, 60, 62, 230, 339
 ontological/linguistic relativity 55
operator 2, 73, 100, 121, 145–7, 173, 176, 178, 212, 216–19, 221, 225–9, 232–9, 251–
2, 313–16, 322, 332–4, 341–2
 NP 100–21, 147–68, 178–94, 216–38
 clause 216–38, 252
optional 32, 42, 45, 51, 71, 77, 102, 105–6, 108, 110–
11, 115, 119, 141, 154, 182, 186, 312, 340, 348
(p.396) ordering principle 2, 5, 99, 239–335, 337–8, 344, 348–9
 Heaviness Serialization Principle (Hawkins) 284
 LIPOC (Dik) 249, 275, 284, 293–6
 Mobility Principle (Hawkins) 284
 placement relative to the boundary of a domain 250–1
 placement relative to a constituent in a domain 251
 placement relative to a domain 249
 Principle of Centripetal Organization (Dik) 251
 Principle of Domain Integrity 239–40, 250, 253–60, 265, 270, 274, 284, 293–
4, 296, 298, 309, 311, 314, 324–7, 332–5, 338, 344, 348
 Principle of Functional Stability (Dik) 249
 Principle of Head Proximity 99, 203, 251–2, 254, 260, 261–312, 314, 316, 322–
5, 327, 332–5, 338, 344–6, 348–9
 Principle of Iconic Ordering (Dik) 248
 Principle of Increasing Complexity (Dik) 249, 252, 275, 294
 Principle of Matrix Disambiguation (Hawkins) 295
 Principle of Pragmatic Highlighting (Dik) 250, 252, 267
 Principle of Promotion Attachment (Hawkins) 295

Principle of Promotion Construction (Hawkins) 295
Principle of Scope 247–8, 254, 260, 309, 314–35, 338, 344, 346, 348
principles with different kinds of reference points 251–2
Relator Principle (Dik) 251, 281, 293, 295–6
parallels NP/predication 2, 5, 13, 22, 56, 57–9, 104, 213–38, 340, 343
participant coding 70–3, 240–1
participle 17–18, 123–4, 137, 139, 168, 186, 190, 208, 310, 326
particle 15, 72, 110, 115, 120, 123, 134–
5, 139, 172, 186, 199, 210, 270, 274, 279, 283, 298, 317, 321, 329, 345
parts-of-speech 4, 8, 10–18, 29, 92, 120, 122–3, 129, 140–1, 145, 172, 190, 285–
9, 302, 311, 334, 341, 348
rigid/flexible distinction 10–19, 28, 42–4, 106, 122–3, 133, 266, 288–
9, 308, 339, 347
person 12, 14, 17, 21, 31–2, 38, 44, 51, 55, 64, 72, 75, 79, 82–5, 88, 106, 108, 112–
14, 119–20, 130, 136–7, 139, 152, 157, 164, 166, 169–70, 174, 186, 190–1, 193–
4, 197, 202, 206, 210, 232, 235, 241, 246, 286–7
person-number-gender marker 31, 82–4, 123, 127, 170, 194
phasal adverbial 236
phonology 23, 25, 60, 69, 96–9, 226, 253, 260, 339
epithetic vowel 102
long/short vowel distinction 80, 96, 115, 188
stress 135, 185, 191, 250, 260; *see also* emphasis
tone 32, 60, 98–9, 116, 126, 130, 151, 162, 245
vowel harmony 43, 80, 164
phrasal/non-phrasal categories (Dryer) 278–91
plural marking, *see* number plural word 32, 119, 279, 285–7, 298
polarity 81, 85–6
politeness 23, 26, 64, 72
positional verb 63–7, 69, 86, 186
possession 2, 13, 21, 24, 43–4, 51, 62, 64, 66, 72–4, 81, 86–91, 94, 98–9, 115, 165–
6, 168, 173–7, 181–2, 186, 194–205, 219, 222, 224–5, 254, 257, 261, 263–4, 266–
7, 271–2, 276, 278, 283–4, 290–1, 299, 300–8, 314, 316, 319, 322, 325, 329, 333, 336–
7, 342, 345–6, 348–9
alienable/inalienable distinction 60–99 (esp. 87–91), 197–9, 271, 290–1, 299–
303, 305–7, 309–11
bound possessor pronoun 197–9, 271, 306–7
doubling of possessor NP 203–5
doubling of possessor pronoun 199–200
possessive classifier 90
possessive pronoun, *see* possessor pronoun
possessor ascension 90
possessor deletion 89
possessor NP 200–5, 306–7, 309–11
possessor pronoun 194–203, 305–7
postfield 251–2, 262, 332
postnominal 20–2, 32, 124, 126–7, 131, 140, 160–2, 164–5, 181, 183, 199–201, 203–
4, 207–8, 210, 261, 266–7, 275–6, 295, 302, 308–11, 318–19, 327, 333, 335, 340
postposition, *see* relator
pragmatic factor 3, 8, 11, 22, 25, 37, 42, 58, 108, 137, 155, 160–1, 174, 226, 229–
31, 244–5, 250, 252, 257, 267, 273, 275, 320, 325, 335, 338

(p.397) pragmatic function 8, 25, 226
see also focality; topicality
afterthought 205, 209, 243, 245
tail 317
theme 20
predicate noun 28, 56–7, 93
predicate satellite 228, 237
predicate variable 226–38
predication 2, 5, 11, 133, 214, 218, 223, 225, 232, 235, 237–9, 281, 312, 321, 344, 347–8
layer 213–38
tensed/non-tensed 233
prefield 20, 251–2, 262, 324, 332
prenominal 20–1, 33, 131, 139–41, 160–1, 165, 167, 185, 197, 199–201, 204, 208, 255, 266, 275, 295, 302, 304, 309–11, 318–19, 340
preposition, *see* relator
presupposition 176, 230
preverb 245
pronoun 2, 12, 17, 19, 22, 25, 31–2, 35, 38–9, 43, 51, 71, 73, 78–9, 83, 87–9, 90, 92–3, 97, 99, 105, 108, 111, 113, 119, 126, 130, 153–5, 165, 169–70, 173–5, 178–9, 180–2, 185, 188–9, 191, 194, 196–203, 206, 219, 224, 241–2, 245–6, 249, 254, 256–8, 262, 266–7, 271–2, 285–7, 294, 301–2, 305–7, 312, 316, 318–19, 325, 327, 340, 342, 348
proper name 19, 94, 202, 262
proposition 11, 13, 19, 226, 232, 254, 348
propositional modifier 226
question particle 283
qualifying modifier in NP 17, 23, 34, 100–45, 173, 175, 208, 212–13, 218–29, 237–9, 313–23, 332, 336, 340–1, 343, 346–7
qualifying operator 100–21, 173, 212, 218–29, 313–23; *see also* nominal aspect
qualifying satellite 121–45, 173, 175, 212, 216, 218–29, 298, 313–14, 318–23, 332; *see also* adjective
quantifying modifier in NP 34, 146–72, 212, 218–29, 313–35, 337, 341–2, 346–7
quantifying operator 146–68, 173, 218–29, 313–16, 323–7, 341; *see also* number
quantifying satellite 168–72, 173, 218–29, 250, 313–16, 323–7, 332, 341; *see also* numeral noun/verb
quantifier 2, 40, 42, 109, 114–16, 153–6, 165, 258, 272, 286, 296, 319, 329, 337, 341
quantifier float 258
rational/non-rational distinction 36, 38 *see also* human/non-human distinction
realis/irrealis distinction, *see* actual/non-actual
reference 1, 8, 10, 27, 44, 59, 183–94, 226–38, 231
actual referent 27
anaphoric reference 229, 232
anchoring, *see* grounding
discourse referent 27, 28, 44, 55, 231, 234
generic reference 19, 46, 50, 78, 93, 102, 174, 183, 235
grounding 175–7, 232–4, 291, 342
identifiability 173–212 (esp. 175–8), 209, 229–31, 291
immediate referent 28, 55, 336
intended referent 27, 219, 228
reference point 178, 214–15

referent file 177, 228
 referent variable 226–38
 specific/non-specific distinction 22, 78, 89, 92–6, 114, 120, 155, 165, 169, 174, 182–3, 185–6, 227, 235–6
Relational Grammar 90
 relative clause 2, 100, 132, 134–5, 145, 165, 169, 173, 175, 177, 194–6, 203, 205–10, 223, 225–6, 264, 274, 276–7, 279, 283–4, 287, 290–1, 298–303, 307, 309–11, 314, 316, 318, 325, 332–4, 341, 345–6, 348–9
 adjoined 205–6, 302–3, 307
 correlative 205, 207–8, 302, 307
 doubling 208–9
 extraposed 205, 259
 finite/non-finite distinction 208, 310–11
 head-internal 194, 205, 302
 languages without relative clause 100, 207
 as qualifying satellite 17, 134–9, 173, 333–4
 relative clause-adjective order 298–9
 relator 21, 43, 62, 70–2, 166–7, 178, 199, 202–3, 209, 240–1, 245, 251, 281–2, 293, 295–6
 adposition 63, 70, 99, 178, 240, 245, 251, 277, 279, 281, 284, 295–8
 case marker 26, 63, 70–1, 93, 99, 189, 204, 210, 240, 245, 251, 260, 281, 295
 complementizer 178, 279, 281, 295, 298
 conjunction 296
 coordinating conjunction 245, 251, 281, 294–6
 double case 210
 linker 77, 138, 202, 209
 postposition 21, 70, 164, 211, 240, 277, 281–2
(p.398)
 preposition 15, 70, 94, 164, 198–9, 240, 257, 277, 281–2, 292–3, 296–7
 subordinating conjunction 131, 136, 208, 233, 245, 249, 251, 279, 281, 294–6, 298
 restrictor 225–6
 rigid/non-rigid V-final distinction 242–3, 245, 252, 264, 282, 301, 347
 role marking 70–2
Role and Reference Grammar 3–4
 satellite 2, 100, 121, 145–6, 173, 175–8, 194, 196, 205, 209–12, 216–23, 225–33, 236–9, 250, 298, 313–14, 325, 332–4, 341–3
 clause satellite 216–38
 NP satellite 121–45, 168–72, 173, 194–211, 216–38
 scope 217–21, 226–7, 233, 238–9, 247, 254, 260, 309, 313–17, 319, 321–7, 331–3, 335, 338, 344, 346–8
 Sein-correlate 27–8, 44, 55–6, 58–9
 Seinsart 4, 28, 55, 57–9, 60, 62, 79, 101, 146, 156, 228, 338–42, 346–7
 see also nominal subcategories
 Aktionsart 58
 aspect 57–9
 definition 53–6
 selection restriction 68–9
 semantic function 8, 69, 71, 174, 178, 194, 205, 210, 214, 219, 225–6, 298, 312, 343, 348
 agent 67, 69, 178, 226
 benefactive, *see* beneficiary

beneficiary 178, 205, 210
comitative 210
location 71, 79, 86, 90, 93, 173–6, 178, 192, 194, 210, 214, 219, 221–5, 342
patient 66–7, 69, 109
recipient 69
source 194
set noun, *see* noun
Shape (feature) 28, 44, 48, 50–7, 59, 62–4, 66, 73, 78–9, 83, 100, 104–5, 117–18, 121, 141–7, 178, 218, 228, 338–41, 347
similarities NP/predication *see* parallels
similarity 31, 83, 86, 103, 174–5, 213
singleton set 46, 52, 101, 103–5, 141–2, 144, 147, 155, 220, 321, 340, 347
singular object noun, *see* noun
singulative aspect, *see* nominal aspect
sort noun, *see* noun
speech act, *see* illocution
spoken vs. written language 23, 151, 161–2, 166, 204–6, 216, 230, 259–60, 327
Stage I/II/III article 119
stative verb 13, 100, 136, 139, 143, 278, 341
style 26, 73–4, 140, 165, 203, 293
Subject 25, 33, 38, 40, 43, 55, 57, 63, 68, 98, 106, 108, 112–13, 120, 169–70, 192, 241–4, 246, 249, 252, 257–8, 262, 271, 279–81, 298, 324–5, 330
subordination 17, 131, 177, 207, 211, 245, 249, 264–6, 279, 281–2, 291, 294–5, 297–8, 309, 311, 314, 348
see also embedded construction
superlative 318
symmetry 223–4
see also asymmetry; parallels
syntactic function 8, 226
see also Object; Subject
syntax 239–335, 337–8, 344–6, 348–9
basic word order 240–5, 299–309
complex NP 309–11, 332–4
correlation 261, 264, 276–7, 280, 291, 344
free word order 244
Greenberg's *Universal 6* 244
Greenberg's *Universal 18* 304–5
Greenberg's *Universal 20* 273–6
non-rigid V-final languages 240–5
ordering pattern 2, 4–5, 240, 242, 254–5, 264–5, 290, 314, 335, 349
ordering principle, *see* above
OV/VO correlation pairs (Dryer) 278–91
OV/VO distinction 298–311;
 OV 257, 278–80, 282–3, 290, 294, 297, 345–6, 348–9;
 VO 278–83, 290, 297, 345
placement rule 247, 294
P1 position 250, 252
reference point 248, 251–2, 260
SVO 242, 244, 261–2, 276–7, 280–1, 302
V-final language 240–5

V-initial language 240–5
V-second language 240–5
VOS 277, 281
VSO 244, 277, 281
tense 12, 65, 89, 108, 156, 175, 178, 190, 214–18, 224, 232–4, 237, 243, 279, 298, 345
term phrase 10, 14, 19, 236
see also noun phrase
term predicate 225
(p.399) term variable 227
topicality 25, 42, 78, 165, 176–7, 189, 250, 340
transitive/intransitive distinction 38, 95, 106, 108, 139, 236, 244, 262, 330
transnumerality 32, 36, 38–44, 45–50, 100–21, 114–15, 117, 121, 141–6, 219, 320–1, 338–40, 347
underlying structure 2–5, 13, 27, 213, 216, 220, 222–6, 231, 234, 237–9, 248, 253, 280, 293, 313–15, 332, 337, 340–3, 347–8
'underspecification' of noun 49, 56, 339
universal 4, 11, 54, 141, 143, 145, 240, 244, 247, 254, 258, 261, 273–4, 276–7, 281, 284, 295, 301, 304, 308, 332, 337, 340, 344
variables 226–38
variation 1, 3, 5, 26, 43, 73–4, 76, 86, 91, 102, 188, 199, 239, 244, 274–5, 296, 299, 326, 328, 334
verb/noun problem 13
verb patterner (Dryer) 279–91
verbal aspect 12, 57–9, 101, 103–5, 156, 214–18, 244, 340, 347
 egressive 57–8, 103–4, 216
 imperfective 36, 57–8, 103–4, 216
 ingressive 57–8, 103–4, 215–17
 iterative 214–15, 217–18, 224
 perfect 132, 156, 214–15
 perfective 57–8, 103–4, 216
 progressive 65, 103, 130
 prospective 214
 repetitive 214
 retrospective, *see* perfect
 semelfactive 214, 217–18, 224
 stative 12
voice
 anti-passive 10
 passive 44
word classes, *see* parts-of-speech
zero 38, 109, 113, 242 (**p.400**)

(p.401) Index of Languages

- Abaza 323
Abipon 290
Abkhaz 6, 9, 18, 30–1, 33–4, 39, 79, 93, 106–7, 113, 123, 129, 140, 142, 148–50, 158–61, 164, 167, 179, 182–3, 186, 192, 195, 197, 208–10, 240, 244–6, 248, 300, 304, 310, 322–4, 327, 334
Adyghe 310, 323
Afar 97
Agaw 118
Aghem 273
Ainu 77
Akkadian 125
Alamblak 6, 9, 18, 24, 26, 30–2, 78, 80, 82–3, 123, 129, 140, 142, 148, 158, 160–1, 170, 179, 181, 186, 188, 193–5, 208, 240, 300, 304, 324, 328, 331
Albanian 310
Amharic 267
Anejom 120
Apache 68
Arabic 275
 Egyptian Arabic 182
Arapesh 9, 97, 99
Arbore 326
Armenian, East 310
Asmat 63, 65, 118–19, 245, 323
Assamese 166
Assyrian 310
Athabascan 182
Awngi, *see* Agaw
Babungo 7, 9, 18, 24, 30, 34–5, 80–2, 129–30, 133, 142, 148–51, 158, 167, 170–2, 177, 179, 181, 186, 188, 195, 199, 209–10, 222, 230, 240, 245, 247, 258, 267, 270–1, 273–4, 276, 300, 302, 304, 324–5, 327, 335
Bambara 7, 9, 18, 30–2, 72, 119, 129–30, 133, 142–3, 148, 151, 158, 170–1, 179, 181–2, 186, 195, 207, 240, 247, 300, 302, 304, 307, 324, 327–8, 331
Banda 289
Bargam 89
Bari 289
Bashkir 310
Basque 6, 9, 18, 24, 30, 38–9, 70, 79, 101, 107–8, 113, 123–4, 129, 140, 142, 148, 158–61, 177, 179, 186, 189, 195, 208–10, 240, 244, 300, 304, 310, 324, 327–8, 331
Beja 275
Bengali 4, 166, 182
Berbice Dutch Creole 7, 9, 18, 30, 38–9, 106–8, 123–4, 142, 148, 152, 158, 179, 182, 186, 188, 195, 199–201, 203, 240, 245, 300, 302, 304, 324, 328, 331
Breton 310
Brôu 4, 25
Bukiyip 6, 9, 18, 20, 30, 35, 81–2, 123–4, 129, 140, 142, 148–51, 158, 160, 179, 182, 186, 195, 199–201, 203, 233, 240, 266–7, 273, 300, 304, 328–9
Bulgarian 182, 310

Burmese 7, 9, 18, 21, 30, 41, 48–9, 75–7, 119, 133, 135–6, 141–2, 144, 148, 153–5, 158, 162, 164, 179, 186, 195, 234, 240, 247, 278, 300, 304, 324, 327
Burushaski 6, 9, 18, 30, 35, 88, 107–9, 120, 123–4, 129, 142, 148–50, 152, 158, 164, 179, 186, 192, 195, 197, 240, 300, 304, 307, 324, 327–8, 331, 334
Caddo 69
Cambodian 289
Cayapo 290
Cayuga 6, 9, 12–13, 17–18, 22, 60, 91–2, 133, 142, 148, 158, 179, 186, 195, 240, 245, 301, 303, 311
Cebuano 77
Chamicuro 224
Chamorro 119, 287
Chinese (Mandarin Chinese) 4, 7, 9, 18, 30, 41, 75–6, 78, 103, 107, 115, 117, 133–5, 141–2, 144, 147–8, 153–4, 156, 158, 162, 164–6, 179, 184–6, 188, 191, 195, 241, 244–5, 247–8, 289, 299–300, 304, 324, 327, 332
Chinook 84, 193
Chipewyan 68
Chräu 4, 166, 245
(p.402) Chukchi 6, 9, 18, 26, 30–1, 33, 36, 38–9, 107, 109, 122, 129, 132–3, 140, 142, 148, 158–60, 179–81, 186, 189, 196, 218, 240, 244, 300–1, 304, 324, 327, 329–30
Chumash 153
Chuvash 103, 310
Coos 245
Cuna 103, 118–20, 157, 163, 168
Dagomba 289
Dakota 290
Dargwa 310
Dasenech 131, 275
Deiga 324
Duala 289–90
Dutch 6, 9, 18, 29–30, 35, 39, 44–6, 50, 52, 57, 61, 63–5, 80, 90, 101–3, 108, 119, 123–4, 129, 142, 147–9, 153, 158, 179, 186, 189–90, 195, 199–201, 203, 208–10, 230, 241–2, 245–6, 249–50, 259, 300–2, 304, 310, 324, 328–9, 331
Dyirbal 234
Easter Island, *see* Rapanui
Efate 182
Ekagi 290
Enga 63, 65
Erzya 21
Eskimo 71
 see also West Greenlandic
Etruscan 6, 8–9, 18, 142, 301, 303
Ewe 119, 175
Fali 94
Fijian 90, 93, 95, 103, 185
 Boumaa Fijian 24, 44, 91, 103, 152, 166, 169, 219, 285, 332
Finnish 310, 331
French 81, 109, 180, 331
Fula 61

Fulani 289
Gabra 274
Galela 6, 9, 17–18, 30, 38–9, 70, 79, 93, 106–7, 110, 120, 133, 137, 142–3, 148, 158–9, 163, 166, 169, 179, 182, 184, 186, 195, 241, 245, 300, 304, 316, 324, 327, 332
Galla, *see* Oromo
Garawa 277
Gbeya 289
Georgian 7, 9, 18, 21–2, 24, 30, 38–9, 105, 107, 109, 123–5, 129, 140, 142, 148, 158, 179, 183, 186, 188, 195, 203, 208, 241, 300, 304, 310, 324, 328
German 59, 61
Gilyak, *see* Nivkh
Godoberi 310
Gooniyandi 20, 255
Greek
 Ancient Greek 182
 Modern Greek 81, 166, 259, 310, 335
Guajajara 277
Guaraní 6, 18, 24, 38–9, 120, 123–4, 142, 148, 158, 179, 182, 184, 186, 189, 195, 241, 244, 300, 316, 318–19, 321, 324, 331
Gude 6, 9, 18, 30, 38–40, 92–3, 107, 110, 129–31, 133, 140, 142, 148, 158, 166–8, 179–81, 186, 189, 195, 197–9, 209–10, 241, 244, 247, 300, 302, 304, 307, 324, 327, 332
Gumbainggar 290
Hausa 80, 198, 289
Hawaiian 286
Haya 89
Hidatsa 215, 217
Hindi 207, 331
Hittite 6, 9, 18, 30, 35, 38, 63, 70–1, 107, 111, 113, 123, 125, 129, 140, 142, 148–9, 158, 179, 186, 195, 197, 199, 201, 207, 241, 245, 300, 302, 304, 307, 324, 327–8
Hixkaryana 6, 8–9, 18, 22, 25, 30–1, 33, 38–40, 72, 79, 103, 107, 110, 113, 120, 133, 138, 142–3, 148, 158, 171–2, 179–80, 186, 195–6, 207, 241, 256–7, 286, 290, 300, 304, 312, 316–17, 319, 321, 324, 327, 329
Hmong 25–6, 191
 Hmong Njua 6, 8–9, 18, 30, 38, 51, 75, 107, 115–17, 123, 126, 129, 141–4, 148, 158, 162, 164–5, 179, 184–6, 195, 241, 245, 247, 300, 304, 316, 318–19, 324, 328, 331, 340
Hopi 153
Hungarian 4, 7, 9, 18, 30, 38–40, 79–80, 87, 93, 102, 107, 111, 123–5, 129, 142–3, 148, 158, 164, 179–80, 182, 187, 189, 191, 195, 197, 199, 203–4, 208, 241, 244, 266–7, 273, 300, 304, 310, 324, 328, 331, 335
Hurrian 6, 9, 18, 42–4, 60, 106, 119, 122–3, 129, 142–3, 148, 158–60, 170, 179, 187, 195, 241, 245, 300, 304, 308, 324, 327
(p.403) Icelandic 310
Igbo 79, 289
Ika 6, 9, 18, 30, 38–40, 63–4, 107, 111, 123, 125, 129, 142, 148, 152, 158, 160–1, 179, 187, 195, 199–200, 203–5, 241, 300, 302, 304, 316, 318–19, 324, 327–8
Indonesian 331
Ingush 310
Iraqw 119
Irish 182, 310

Italian 170, 182, 331
Jabem 289
Jacaltec 73, 163, 236, 285–6, 331
Japanese 267
Javanese 26
Jukun 87, 89
Kabardian 323, 331
Kalam 25
Kalkatungu 19–21, 23, 255–6
Kalmyk 186, 310
Kanuri 120
Katu 4, 25
Kayardild 6, 9, 17–18, 30, 38–40, 75, 107, 111, 123, 126, 129, 140, 142, 148, 152, 157–8, 160, 170, 179, 187, 195, 207, 241, 244, 300, 304, 324, 328, 331
Ket 7, 9, 18, 30, 36, 70–1, 80–1, 107, 112, 116, 123, 126, 129, 142, 148–9, 158, 179, 187, 195, 241, 244, 300–1, 304, 324, 328, 331
Kharia 153
Khasi 84, 193
Kikongo 98–9
Kikuyu 273, 275
Kilivila 163, 270, 277
Kinyarwanda 98
Kirmanji 310
Kirundi 289
Kisi 7, 9, 18, 30, 34–5, 81–2, 129–31, 133, 142, 148–51, 158, 179, 195, 241–2, 245, 247, 300–1, 304, 328–30
Koasati 6, 9, 18, 26, 30–1, 33, 36, 63–4, 80, 107, 112–13, 118–19, 133, 138, 142–3, 148, 168, 179, 187, 189, 195, 207, 241, 244, 300, 316–17, 324, 327
Kobon 25, 156–7, 233, 290, 312
Kombai 63
Korean 9, 18, 21, 24, 26, 30, 41, 70–2, 75, 103, 107, 133, 135–6, 141–2, 144, 148, 153–4, 158, 160, 162, 164–5, 177, 179, 187–8, 195, 208–9, 233, 241, 243, 247, 300, 304–5, 324, 327, 331
Krongo 7, 9, 18, 24, 30, 33, 36, 70–1, 78, 80–1, 85–6, 118, 133, 138, 142–3, 148–9, 151–2, 158, 167–8, 179, 187, 195, 209–10, 241, 245, 300, 302, 304, 324, 327, 329–30
Kurku 125–6
Kwakiutl 111
Lahu 299
Lango 7, 9, 18, 30, 38–40, 72, 105, 107, 113, 133–4, 142, 148, 158, 172, 179–81, 187, 189, 192, 195, 241, 243, 300, 302, 304, 324, 327, 334
Latin 19, 65, 96
Laz 310
Lenakel 95
Lezgian 310, 334
Lingala 98
Lisu 21
Lithuanian 310
Logoli 274
Loma 182
Lomongo 98

- Luganda 182
 Luo 274
 Lyele 182
 Maa 274
 - Camus 274–5
 - Sampur 274–6
 Maasai 182
 Macedonian 182
 MalakMalak 62
 Malay 165
 Maltese 310
 Mandinka, Gambian 93–4, 120, 323
 Mangarayi 20, 70–1, 103, 153, 160, 170, 189, 256, 290
 Mangbetu 93–4
 Maori 290, 331
 Margi 289
 Masa 102
 Maung 331
 Maya
 - Mocho Maya 236
 - Yucatec Maya 44–7, 49–52, 55, 146
 Menomini 290
 Meroitic 7–9, 18, 142, 301, 303
 Miao, *see* Hmong
 Minor Mlabri 168
 (p.404) Miskito 290
 Mixtec 285–6, 290
 Mohawk 12–13
 Mokilese 77, 90–1, 104
 Monumbo 290
 Motu 290
 Muna 95
 Mundari 18
 Mupun 287
 Nadëb 67
 Nahali 7–9, 18, 31, 38–40, 107, 115, 125–
 6, 129, 142, 148, 153, 158, 179, 187, 195, 241, 245, 300, 304, 324
 Nahuatl 290
 Nama Hottentot 7, 9, 18, 21, 23, 30–2, 45, 78, 80, 82, 84, 103, 123, 126–
 7, 129, 140, 142, 148, 153, 158, 162, 179, 187–8, 193–
 5, 209, 241, 243, 247, 300, 304, 324, 328, 331
 Nandi 172, 274
 Nasioi 6, 9, 18, 24, 30, 36, 73, 107, 113–14, 123, 127–9, 140, 142, 148–
 50, 153, 158, 163, 177, 179–80, 186–7, 189–90, 195, 201–2, 241, 245, 266, 268–
 70, 273, 300, 304, 324, 327–8
 Navajo 68
 Nenets 299, 310
 Newari 4
 Ngalakan 6, 9, 18, 20, 30, 38, 80–1, 107, 113–14, 123, 128–
 9, 140, 142, 148, 152, 158, 170, 179, 187, 195, 201–

- 4, 207, 241, 256, 289, 300, 302, 304, 307, 328–30
- Ngandi** 153, 245
- Ngarinjin** 119
- Ngiti** 7, 9, 18, 30, 37–8, 88–9, 102, 106–7, 113–14, 123, 128–9, 142, 148–51, 158, 160–2, 166–7, 179, 182, 187, 195, 197–9, 209, 241, 258, 288–9, 300, 302, 304, 307, 324, 327–8
- Ngiyambaa** 20, 101, 118, 160, 180, 189, 191–2, 206, 256
- Nguna** 289
- Nivkh** 6, 8–9, 18, 21, 25–6, 30–1, 33, 38–40, 42, 75, 77, 107, 114, 133, 138, 141–2, 148, 153, 159–62, 164–5, 167, 179–80, 187, 195, 241, 245–6, 300, 304–5, 324, 327, 329
- Nkore-Kiga** 274
- Nogai** 310
- Noni** 273
- Nung** 6, 10, 18, 31, 42, 75, 119, 133, 135–7, 141–2, 144, 148, 153–5, 159–62, 164–6, 179, 187, 195, 209, 211, 241, 247–8, 300, 304, 324, 327, 333–4, 349
- Nunggubuyu** 6, 10, 18, 20, 31, 33, 38, 107, 113–14, 133, 138, 142–3, 149, 152–3, 159–60, 170, 180, 187, 189, 195, 199–200, 203–4, 207–8, 241, 244, 256, 300–2, 304, 307, 324, 327
- Onondaga** 66–8, 90–2
- Oromo** 6, 10, 18, 29–30, 38–9, 41, 44–7, 50, 80, 85, 97, 101–3, 105, 107, 118, 129–31, 133, 140, 142, 147, 149, 151–2, 156, 159, 166, 170, 180, 187–8, 195, 201–4, 208–9, 219–20, 241, 245, 270–3, 275, 301, 304–5, 308, 323–8, 331, 345–6
- Boraana** 80, 102, 188, 209, 271, 323, 325, 328
 - Harar** 325
 - Orma** 80, 102, 271–2, 323, 325, 328
 - Waata** 102, 271, 323, 325, 328
 - Wellegga** 118, 131, 272
- Ossetic** 185, 310
- Pacöh** 4, 24–5
- Pala** 290
- Palauan** 233–4
- Panara** 224
- Papago** 88
- Papiamento** 119
- Pilagá** 234
- Pipil** 6, 10, 18, 30, 37, 88, 107, 129, 131–3, 140, 142, 149, 152, 159, 180, 187, 189, 191, 195, 199, 241, 300, 304, 306, 328–31
- Polish** 4, 257–8, 310
- Ponca** 69
- Port Sandwich** 235
- Portuguese** 111, 172
- Quechua** 15
- Imbabura Quechua** 6, 10, 18, 43, 106, 142–3, 149, 159, 180, 187, 195, 205, 209, 211, 241, 245, 259, 301–2, 304, 308, 324, 328, 331
- Rapanui** 103, 236, 331
- Rendille** 274
- Romani** 310
- Romanian** 267, 310

Russian 58, 96, 104, 259, 310
Saho 119
Sahu 184
Sakao 236
(p.405) Saker 89
Salish, Straits 15
Samoan 6, 10, 13–15, 17–18, 21–2, 26, 42, 44, 60, 106, 122, 129, 142–3, 149, 152, 159, 164, 166–7, 169, 172, 180, 182–5, 187, 196–9, 206, 235, 241–2, 283, 290, 300, 304, 307, 324, 327, 332
Santo 182
Sarcee 7, 10, 18, 30, 33, 37, 63–4, 68, 70, 79, 103, 107, 113, 115, 118, 120, 129, 132–3, 135, 142, 149, 152, 159–61, 164, 168, 180, 187, 196, 241, 259, 301, 304, 316, 318, 324, 327
Sardinian 310
Selepet 331
Seri 69
Sesotho 82, 150
Shilluk 119–20
Shipibo 290
Sidamo 170, 326
Sika 290
Somali 113
Spanish 47, 61, 65, 131–2, 267, 310, 331
Sumerian 7, 10, 18, 24, 26, 30, 38–9, 41, 70–1, 103, 107, 113–14, 123, 125, 129, 140–2, 149, 152, 159, 162, 166, 169, 177, 180–1, 187, 196, 201–4, 241, 301, 304–5, 324
Svan 310
Swahili 62, 151, 289
Swedish 4, 310
Takelma 233
Tamil 6, 10, 18, 24, 30, 38, 70–1, 107, 113, 115, 126, 129–30, 133, 142, 149–50, 159–60, 162, 177, 180, 185, 187–8, 193, 196, 208, 241, 245, 267, 270, 272–3, 301, 304, 324, 328, 331, 334–5
Tanna 182
Tarahumara 174
Tati 310
Teso 85
Thai 29, 44–52, 55, 75, 77–8, 289
Tigre 277
Tigrinya 84
Tlingit 103, 118
Toba 71, 234
Tobelo 137, 184
Tongan 15, 18
Trukese 91
Tsou 6, 10, 18, 31, 38–9, 41, 103, 107, 115, 122, 133, 139, 142–3, 149, 152, 159, 180, 187, 196, 208–9, 241, 243, 247, 277, 300–1, 304–5, 307, 324, 327
Tunica 87
Turkana 84, 274
Turkish 4, 6, 10, 18, 42–4, 60, 106, 122, 129, 142–3, 149, 159, 180, 185, 187, 196–7, 199, 203–4, 208, 241, 243, 245, 266–8, 273, 301, 304, 308, 310, 318–

- 22, 324, 328, 331, 333–4, 349
Tuscarora 12–13, 92, 174
Ubykh 323
Udmurt 310, 334
Upper Chehalis 119
Usan 25, 312
Vietnamese 6, 10, 18, 31, 42, 75, 77–8, 133, 135–6, 141–2, 144, 149, 153–5, 159–62, 164–5, 180, 184–5, 187, 191, 196, 241, 247, 285–6, 289, 300, 304, 324, 327
Warlpiri 206, 256
Wambon 6, 10, 18, 20, 30, 34, 38–9, 41, 107, 114, 123, 129, 142, 149, 152, 159–60, 180, 187, 196, 205, 241, 289, 301–2, 304, 324, 328–31
Warao 331
Wardaman 20
Waris 69
Welsh 331
West Greenlandic 6, 9, 18, 24, 31, 37, 133, 139, 142, 149–50, 152–3, 159, 166, 170, 177, 180, 185, 187, 196–7, 199, 209, 211, 233, 241, 248, 300, 304, 324, 326–7
Wolof 289
Yagan 289
Yagua 73, 163, 277
Yanomama 232
Yao 21
Yavapai 290
Yidiny 75
Yimas 20, 97–8
Yoruba 119, 289, 331
Yurok 70
Zande 304, 332 (p.406)

(p.407) Index of Authors

- Abaev 185
Abney 4, 280
Adams 48, 50, 59, 76, 165–6
Aikhenvald 60
Allan 62, 68, 71, 75–6, 101, 254
Allen 67, 312
Ameka 62
Anceaux 95
Anderson, J.M. 103
Anderson, S.R. 101, 118, 121, 151–2, 174, 178, 185, 214
Andrews 8
Andrzejewski 45
Anward 10
Archangeli 239
Aristar 194
Aristotle 4, 213, 215, 229, 336, 347
Arnott 61
Aronoff 23
Asher 18, 24, 38, 71, 130, 162, 177, 185, 188, 193, 208, 245, 270, 272, 334
Austerlitz 41, 77, 153
Auwera 4, 236
Bach 4, 177
Backhouse 10
Bahr 88
Bakker 5, 243, 291
Barnes 245
Barron 63, 68, 165, 186
Bartsch 156, 220
Beck 10
Becker 76–7
Beekes 103
Beffa 138
Behaghel 249, 253, 265, 294, 337, 344
Behrens 93
Bender 85, 131, 245, 270–2, 274–5, 323, 325, 345
Benton 76
Benveniste 4
Benzing 101, 103, 186
Berger 35, 124, 152, 186, 192, 197
Berman 207
Bever 254
Bhat 10
Bickerton 23
Biermann 101
Bird 72
Birk 62
Bisang 45, 78, 116, 141, 166
Blake 19–20, 90, 205, 246, 255–6, 307

Blankenship 68, 76
Boas 45, 63, 111, 190
Bodine 26
Bogoras 39, 109, 181, 189
Bolinger 140
Bonfante 8
Bonvillain 13
Booij 121
Borg 101
Börjars 4
Bot Ba Njock 274
Bouma 258
Bradley 286
Brauner 18, 32, 119, 130, 171, 181–2
Braunmüller 23
Brinton 101
Broschart 15, 18
Bruce 24, 26, 31–2, 83–4, 123, 140, 161, 170, 181, 188, 193–4, 208, 328
Bryan 85, 94, 102, 119
Bulteau 285
Bunt 48
Burling 50, 76
Bybee 65, 103–4, 121, 232–3, 245
Campbell, G. 8
Campbell, L. 8, 18, 37, 88, 131–2, 152, 191, 199, 306, 330
Capell 26
Carter 68
Castrén 36, 112
Chafe 233
Chao 103, 134, 154, 165
Chappell 62, 86, 135
Childs 18, 35, 82, 92, 131, 150, 242, 245, 330
Christensen 344
(p.408) Christie 174
Chung 103–4, 232–4
Churchward, C.M. 15, 18, 44, 103
Churchward, S. 14
Clark, E.V. 76, 174–5, 178, 218, 265–6, 296, 308
Clark, H.H. 218, 265–6, 296, 308
Claudi 175, 219, 223
Cole 43, 205, 211, 245, 259
Comrie 5, 19, 21, 26, 39, 42, 58, 70–
1, 77, 81, 103, 109, 113, 117, 132, 156, 161, 164, 176, 186, 196, 205, 207, 214, 232, 234, 244
5
Conklin 76
Connolly 248
Conrad 20, 35, 82, 124, 141, 150, 182, 199, 201, 266–7, 329
Cook 18, 37, 56, 64, 68, 79, 103, 115, 118, 120, 132–3, 161, 164, 168, 259, 316, 318
Copeland 174
Corbett 21, 60, 73, 84–5, 96–7, 155, 168, 170–1

Cornyn 155
Costello 4, 25
Court 21
Craig 4, 41, 60, 73, 163, 236, 286, 336
Creider, C. 82, 172, 274
Creider, J. Tapsubei 172, 274
Croft 3, 10–11, 21, 86, 153, 196, 228, 239, 243, 253
Crowley 95–7, 120, 183, 235
Cusic 101
Daeleman 98
Dahl 93, 101, 104
Davidson 68
Davies 25, 156, 233, 312
de Groot 71, 103, 197, 244
de Reuse 56
de Vries 20, 27, 34, 41, 63, 114, 129, 152, 160, 205, 289, 331
DeBose 119
DeLancey 239
Delsing 4
Demers 15
Demuth 82
Denny 68, 76, 82
Deny 18
Derbyshire 4, 21–2, 25, 33, 40, 72, 80, 103, 110–
11, 120, 138, 163, 171, 180, 196, 207, 256, 312, 317, 321, 329
Devitt 65
Diessel 174, 178
Dik 2–4, 8, 13, 19, 23, 26–8, 56, 58–9, 70, 76, 101, 103, 151, 155, 174–
5, 178, 189, 205, 212–14, 216, 225–8, 231–3, 235, 237, 239–40, 243–6, 248–
51, 260, 262, 264, 275, 280–1, 283–4, 292–5, 307, 321, 335, 337–8
Dimmendaal 84
Dixon 24, 44–5, 61, 68, 70, 73–7, 80, 91, 95, 99, 103, 111, 153, 169–70, 219, 285
Dmitrijev 267
Donaldson 20, 101, 118, 153, 160, 180, 191–2, 206, 256
Donnellan 235
Dotter 245
Downing 25, 205, 207
Drabbe 63, 118
Dressler 152
Drohla 111
Drossard 26, 38, 40–2, 48, 114, 138, 153, 161, 181, 246, 329
Dryer 24, 32, 119, 185, 188, 240, 243–4, 263–4, 266, 277–87, 290–1, 297–9, 302, 304–
5, 308, 345–6
Du Feu 103, 236
Dunn 26
Durie 152
Edwards 245
Elliot 233
Emeneau 136, 154–5
Erbaugh 76

Evans 17, 40, 62, 75, 111, 126, 140, 152, 157, 168, 170, 180, 207, 244, 328
Ewert 4
Fähnrich 39, 109, 125, 203, 241
Farsi 218
Ferguson 102
Fleischman 223
Fodor 60
Foley 4, 20–1, 25–6, 41, 55, 63, 69, 73, 75, 77, 97, 163, 177, 228, 340
Fortescue 24, 37, 139, 166, 170, 177, 180, 185, 197, 211, 227, 233, 326
Fortune 97
Fox 90, 176, 205, 245, 307
Frajzyngier 31, 38, 110, 116, 119–20, 152, 178, 287
Frank 40, 64, 111, 125, 152, 161, 200, 204–5, 319
Fraser 168
Frazier 263, 266
(p.409) Frege 229
Frei 214
Freiberger Wilson 42, 135–7, 154–5, 161, 211, 333
Friedrich, J. 36, 111, 125, 199, 201, 207
Friedrich, P. 54, 79
Fries 218
Fuller 25
Gabler 55
Gandour 48
Gangopadhyay 4
Gao 4, 134
Gardiner 67
Garrett 71, 111, 125, 245
Gasparov 259
Geerts 57, 201, 250
Gentner 55
Gernot 43
Gerstner-Link 50, 93
Gil 4, 48
Gildea 174, 178, 224
Gilligan 242
Givón 25, 177, 186, 207, 211, 253, 262
Goodman 51–2
Goral 21, 60, 76, 165–6
Gorbet 25
Gore 332
Gragg 118, 131, 209, 272
Greenberg 21, 34, 45, 50, 60, 73, 76–7, 80, 84, 92–7, 101, 118–20, 152–3, 156, 160–6, 170–1, 174, 178, 180, 182, 185, 188, 191, 193, 214, 235, 239–40, 244, 253–5, 261, 273–8, 284, 302, 304, 326, 340–1, 344
Gregores 24, 33, 40, 120, 125, 184, 190, 244, 318–19
Grice 24–5
Grierson 8
Grimes 8
Grinevald 4, 69, 73, 336, 340

Grönbech 43
Grosjean 254
Gruzdeva 40, 77, 114, 138, 153, 161, 164, 167, 245–6
Gundel 243
Guy 236
Gvozdanović 156
Haas 26, 68, 87
Hagège 274
Hagman 21, 31–2, 84, 127, 162, 188, 193, 243
Haiman 23, 25–6, 74, 140, 253
Hale 170, 205–6, 256
Hannay 176
Hansen 4
Harder 234
Harriehausen 61, 115–17, 126, 142, 184, 245, 247, 319
Harris 93, 105, 109, 208
Harrison 87, 91, 95
Harvey 75, 336
Hashimoto 332
Haspelmath 100, 223, 310, 334
Haugen 23
Hawkins, E.A. 286
Hawkins, J. 4, 140, 208, 239–40, 245, 261–3, 273–7, 284, 291–9, 305, 307, 331–2, 335, 345
Haycock 8
Hayward 85, 97, 326
Heath 20, 33, 101, 114, 138, 152, 160, 180, 189, 200, 204, 207, 244–5, 256
Heim 177
Heine 4, 84, 97, 156, 175, 178, 185, 219, 223, 274–5
Held 207
Hengeveld 10–12, 16–17, 28–9, 61, 63, 122–3, 132, 139–41, 216, 226–8, 231–3, 236–7, 287–8, 302, 308, 321, 347
Herring 23
Hetzron 85, 118, 170, 218, 267, 326
Hewitt 34, 79, 106, 123, 140, 150, 159, 161, 164, 167, 183, 192, 197, 208–10, 244, 246, 322, 334
Hill 26
Himmelmann 174, 185
Hindle 23
Hjelmslev 178
Hock 208, 245
Hoffmann 18
Hoijer 68
Holm 4
Holmberg 124
Holmer 103, 119–20, 157, 163, 168
Holton 137, 184
Honey 136
Hope 21
Hopper 11, 103, 165

Hoskison 18, 38, 40, 93–4, 110, 130–1, 140, 167–8, 180–1, 189, 197–8, 210, 244, 332
Hovdhaugen 14, 22, 42, 152, 164, 167, 169, 183–4, 198–9, 206, 235, 283
(p.410) Huang 135
Hudson, G. 85
Hudson, R.A. 275
Hundius 29, 45–50, 78
Hurd, C. 73
Hurd, Ph. 73
Hurford 156, 170
Hyman 89, 273
Hymes 23
Ide 26
Iljic 147, 154
Imai 55
Jackendoff 101, 280
Jaisser 165
Jakobson 26, 138, 175, 329
Jelinek 15
Jespersen 117
Johanson 233
Johnson 174, 178, 223, 253
Jones 4, 21, 45, 77–8, 165–6
Joseph 81, 166, 259
Josephs 234
Jurafsky 118–19
Justus 207
Kahr 245
Kahrel 245
Kamp 176
Kämpfe 109, 132, 160, 196, 218
Kang 154
Kant 228
Karttunen 27, 55
Kastenholz 32, 143, 207
Kaufman 23
Kazenin 310
Keenan 174, 176, 178, 185, 205, 207, 214, 239
Keizer 19, 227
Kiefer 87
Kim, N.K. 188, 243
Kim, Y.H. 154
Kimball, G.D. 26, 33, 64, 80, 112, 118, 138, 168, 190, 207, 244, 317–18
Kimball, J.P. 296
Kinkade 119
Kirk 113
Kirsner 74
Kiyomi 152
Klein 234
Kölver 4, 29, 45–50, 78, 155
Köpcke 61

Koptjevskaja-Tamm 19, 38
Kornfilt 204
Kouwenberg 39, 108, 124, 152, 188, 199, 201, 203, 245
Krámský 185, 188, 267–8
Krauss 68
Krejnović 161
Krifka 101
Kuh 154
Kuhn 101, 121
Kuiper 8, 40, 115, 125
Kuno 205
Kuryłowicz 117
Kutsch Lojenga 37, 88–9, 114, 128, 151, 161–2, 167, 182, 198, 258, 288–9
Labov 23
Lacko 4
Lakoff 61, 174, 178, 223, 253
Lambrecht 226
Landar 68
Landsberg 253
Lang 63
Langacker 101, 119, 228
Langendoen 239
LaPolla 4, 101, 188, 213, 254
Lappin 4
Larochette 94
Laury 186
Lê Văn Lý 286
Lee, H.H.B. 18, 21, 24, 33, 41, 71–2, 136, 162, 164, 177, 209, 243
Lee, H.-G. 154
Lee, H.S. 136
Lee, Y.-S. 103, 136, 144, 154
Leech 101
Lefebvre 43
Lehman 76, 163
Lehmann, C. 19, 21, 31, 60, 73, 78, 119, 140, 152, 166, 177, 180, 205, 207, 276, 287, 311
Lehmann, W.P. 278
Leinonen 259
Levinson 26, 178, 214, 223
Lewis 42–4, 122, 197, 204, 208, 245, 268
Li 41, 76, 78, 126, 134–5, 144, 154–5, 165, 184–5, 188, 244–5
Lichtenberk 90
Limburg 71, 196–7, 245
Lindsley 254
Linell 23
(p.411) Ljung 86
Löbel 78, 135–6, 185, 191
Lönnqvist 259
Lorimer 35, 108, 120, 124, 150, 164, 187, 192, 197, 334
Lotz 267
Lucy 46–7, 49–51, 55, 101, 146

Lukas 119
Luong 155
Luraghi 36, 111, 125
Lyons, C. 185
Lyons, J. 10, 19, 50, 76, 103, 174–5, 178, 214, 232–3, 340
McCawley 4, 27, 135
McCool 180
MacDonald 245
McGregor 20, 23, 62, 86, 245, 255
Mackenzie 19, 87, 211, 231, 245
Malchukov 168
Mallinson 205, 259, 307
Manzelli 182, 196
Martin, L. 236
Martin, S.E. 103, 136, 144, 154
Matthews, G.H. 217
Matthews, P. 22, 215, 256
Mattissen 26, 38, 40–2, 114, 138, 153, 161, 181, 246, 329
Maxwell 205
Meillet 4, 70, 117
Meinhof 84–5, 193
Meisterfeld 101
Menges 42, 117, 123, 186
Menninger 156
Merlan 20, 38, 63, 71, 81, 103, 114, 128, 140, 152–3, 160, 170, 189, 202, 204, 207, 256, 289, 330
Miao Language Team 116, 185
Miller 4, 25
Mithun 13, 22, 26, 44–5, 56, 60, 69–70, 110, 112, 114, 133, 152, 186, 244–5, 311
Mithun Williams 13, 174
Moore 266
Moravcsik 4, 40, 80, 105, 111, 125, 143, 152, 177, 180, 182, 185, 188, 191, 204, 220, 247, 2
Moreno 170
Moreton 93
Mosel 14, 22, 42, 152, 164, 167, 169, 183–4, 198–9, 206, 235, 283
Mourelatos 101
Mous 119
Mpaayei 152
Mühlhäusler 23
Muysken 43
Nedergaard Thomsen 227
Newman 38, 44, 80, 94, 102, 120, 152
Ney 218
Nguyễn D.H. 42, 136, 161, 164, 185, 285–6
Nguyễn D.L. 155
Nichols 233, 247, 310
Noonan 38, 40, 72, 105, 113, 134, 172, 178, 181, 189, 192, 243, 334
Norman 154, 165
Nuyts 3, 5
O'Dowd 223

O'Neill 23
Ochs 26
Okell 41, 48–9, 135–6, 154–5
Owens 170, 325–6
Palmer 11, 232–3
Panfilov 40, 114, 161, 165, 246
Parker 224
Pawley 25
Payne, D.L. 21, 73, 163
Payne, J. 4
Perkins 23, 25, 176, 214
Pesot 35, 108
Philippaki-Warburton 81, 166, 259
Philips 26
Piau 63
Plank 4, 10, 81, 182, 210
Poppe 123
Prasithrathsint 10
Prince 175
Pullum 4
Pustet 10
Quesada 31
Quine 55
Quirk 211
Rabel-Heymann 84
Ramstedt 154
Rankin 63, 69, 186
Rapaport 27, 56
Ratliff 116, 126
Rauh 178, 214
(p.412) Rausch 24, 36, 73, 113–14, 127–8, 140, 177, 190, 202, 268–9
Rebuschi 108
Reesink 25, 312
Reh 24, 36, 71, 85–6, 118, 138, 143, 151–2, 167–8, 210, 245, 324
Reid 75, 336
Reyle 176
Rijkhoff 4–5, 12, 25, 39, 53, 72, 101, 103, 105, 117, 124, 145, 163, 171, 175, 177, 209, 250–1, 261–4, 274, 286, 291, 293–4, 296, 298–9, 309, 311–12, 340, 343–4
Rijksbaron 4, 213
Rischel 168
Risselada 218
Ross 168
Rowlands 94, 120, 323
Ruhlen 5–7, 85
Sackmann 135
Saltarelli 24, 39, 70, 79, 101, 108, 124, 140, 160, 177, 189, 208, 210, 244
Sanches 29, 115, 153
Sanders 3
Sands 74, 92
Sapir 111, 233

Sasse 10, 12–13, 91, 98, 103, 131, 274
Saul 42, 135–7, 154–5, 161, 211, 333–4
Schachter 10, 15, 154
Schadeberg 85
Schroeder 4, 42–3, 185, 243, 319–21, 333
Schwartz 205, 276
Searle 11, 231
Seiler, H. 4, 60, 86, 89–91, 156, 196, 218, 220, 275
Seiler, W. 69
Sellner 76
Senft 60, 163, 270, 336
Seok 154
Serzisko 63, 81, 85, 165, 186
Seuren 101
Sichelschmidt 218
Siewierska 4, 71, 159, 186, 209, 240, 243, 245–7, 257, 307
Silverstein 71
Simpson 206, 256
Sinha 18
Skorik 132, 196
Slobin 29, 115, 153, 254
Smith-Stark 70, 112
Sohn 154
Song 154, 243
Speiser 42–3, 123, 160, 170
Spruit 159, 322
Srivastav 207
Stampe 156
Starosta 139, 301
Stassen 232, 289–90
Steele 239
Stengaard 65
Stoltenberg 8
Stroomer 29, 41, 45–7, 80, 102–
3, 105, 118, 147, 149, 152, 166, 170, 188, 202, 204, 209, 219, 271, 305, 323, 325–6, 328
Suárez 24, 33, 116, 120, 184, 244, 318–19
Sun 78, 165, 185, 191
Swadesh 111
Swanton 103
Szabolcsi 4
Szakos 41, 103, 115, 139, 152, 209, 243, 277
T’sou 50, 76
Tallerman 333
Taylor 274
Tcheckhoff 16
Tedeschi 103
Teferra 26
Tesnière 140, 276
Testelec 22, 24, 140, 188, 203, 208, 241, 323
Thelin 103

Thomas 4
Thomason 23
Thompson, L.C. 136
Thompson, S.A. 10–11, 19, 41, 76, 78, 134–5, 144, 154–5, 165, 176, 184–5, 188, 244–5
Thomsen 24, 26, 38, 41, 71, 103, 114, 129, 141, 152, 162, 169, 177, 181, 202, 204, 227, 305
Tiffou 35, 108
Timberlake 103–4, 232–4
Tomlin 239, 245, 260, 280
Tommola 103
Tompa 267
Topping 287
Tosco 277
Traugott 178
Trudgill 23
Tsunoda 245
Tucker 85, 94, 102, 119, 152, 274, 332
Tuite 117
Tung 152
(p.413) Uhlířová 257
Ultan 87, 89, 174, 180, 196, 232, 306–7
Underhill 42–4
Unseth 102
Unterbeck 55
Utas 233
Valois 4
Valstar 12, 347
van Baar 236
van Baarda 17, 39, 70, 106, 110, 137, 163, 169, 184, 245, 316
van de Kerke 43
van den Berg 95
van der Auwera 4, 236
van der Sandt 176
van Langendonck 253
van Marle 121
van Schaaik 43, 268
Van Valin 3–4, 101, 188, 213, 254
Vendler 19
Vennemann 220, 278
Verkuyl 103
Vet 227
Vidal 234
Vilkuna 21, 258, 311
Visser 184
Vogel 5
Volodin 109, 132, 160, 196, 218
von Garnier 104
von Heusinger 93
Vonen 14
Voorhoeve 118, 184, 323
Vu Duy-Tu' 155, 285

Wang 26, 116, 126
Watkins 63, 69, 186
Watson 4, 24–5
Weber 15
Weir 67
Welmers 79, 84–7, 274
Werner, H. 23, 36, 81, 112, 117, 126, 149, 244
Werner, O. 23
Westermann 119–20
Wetzer 10, 137
Whaley 243, 312
Wheatley 21, 135, 162
Whinnom 23
Whorf 153
Wierzbicka 239
Wilkins 74, 156, 168
Willett 233
Willim 4
Woodbury 66–8, 90–2
Wurm 45, 63, 110
Xrakovskij 152
Yngve 276
Yoo 136
Z'graggen 89
Zaborski 38, 275
Zaenen 103
Zepeda 88
Zimmermann 4
Zipf 25
Ziv 205, 259
Zubin 61 (p.414)

Table of Contents

Title Pages	3
Contents	6
List of Figures	7
List of Tables	8
List of Abbreviations	9
Acknowledgements	15
1 Preliminaries	16
1.1. Introduction	16
1.2. Aims	16
1.3. The theoretical framework	17
1.4. The sample	19
1.5. Nouns, noun phrases, and referents	22
1.5.1. Word classes	24
1.5.2. Nouns	25
1.5.3. Hengeveld's classification of parts-of-speech systems	28
1.5.4. Noun phrases	31
1.5.5. Referents of NPs	37
2 Nominal Subcategories: Seinsarten	43
2.1. Introduction	43
2.2. A preliminary classification	43
2.2.1. Languages with singular object nouns	48
2.2.2. Languages with set nouns	52
2.2.3. Languages with sort nouns	56
2.2.4. Languages with flexible nouns	57
2.3. Towards a comprehensive classification of nominal subcategories	58
2.3.1. Nominal subcategories (reprise)	58
2.4. A new classification of nominal subcategories	64
2.4.1. The feature 'Shape'	64
2.4.2. The feature 'Homogeneity': masses and collectives	65
2.4.3. Defining the nominal subcategories	66
2.5. Nominal subcategories: beyond the noun phrase	68
2.5.1. The semantics of predicate and incorporated nouns	68
2.6. Nominal subcategories as different Seinsarten	69
2.7. Conclusion	71
3 Nouns: Real and Apparent Nominal Subclasses	77
3.1. Introduction	77
3.2. The reflection of real and apparent nominal subclass distinctions outside the NP	79
3.2.1. Positional verbs	79
3.2.2. Classificatory verbs	84
3.2.3. Classificatory elements in the verb stem	85

3.2.4. Predicative numerals	86
3.2.5. Relators and participant coding	86
3.2.6. Honorific nouns and verbs	88
3.3. The reflection of real and apparent nominal subclass distinctions inside the NP	89
3.3.1 Noun classification	90
3.3.2 Noun classes	93
3.4. Conclusion	112
4 Qualifying Modifiers in the Noun Phrase	117
4.1. Introduction	117
4.2. Grammatical expressions of the notion Quality in the noun phrase	117
4.2.1. Aspect	117
4.3. Lexical expressions of the notion Quality in the noun phrase: adjectives	135
4.3.1. Languages with adjectives	137
4.3.2. Languages of intermediate type 3/4: V-N(-A)	144
4.3.3. Languages without a flexible or distinct class of adjectives	148
(p.140) 4.3.4. Doubling of adjectives	155
4.3.5. When can a language have adjectives?	155
4.4. Conclusion	159
5 Quantifying Modifiers in the Noun Phrase	167
5.1. Introduction	167
5.2. Number and cardinality: the grammatical expression of the notion Quantity in the noun phrase	167
5.2.1. Number	167
5.2.2. Cardinality	176
5.3. Numeral phrases: the lexical expression of the notion Quantity in the noun phrase	187
5.3.1. Verbal forms	187
5.3.2. Nominal forms	188
5.4. Conclusion	190
6 Localizing Modifiers in the Noun Phrase	194
6.1. Introduction	194
6.2. Demonstratives and articles: the grammatical expression of the notion Location in the noun phrase	198
6.2.1. Demonstratives	198
6.2.2. Definiteness and indefiniteness markers	204
6.3. Possessive modifiers, adpositional modifiers, and relative clauses: the lexical expression of the notion Location in the NP	213
6.3.1. Attributive possessor pronouns	215
6.3.2. Attributive possessor NPs	219
6.3.3. Relative clauses	224
6.3.4. Other kinds of localizing satellites in the NP	229
6.4. Conclusion	231
7 The Underlying Structure of Noun Phrases	238
7.1. Introduction	238

7.2. Quality, Quantity, and Location in the clause	238
7.3. The layered structure of the clause	240
7.4. The layered structure of the noun phrase	242
7.5. Symmetry between NPs and clauses	247
7.6. Formal representation in Functional Grammar	248
7.6.1. Discourse modifiers	252
7.7. More parallels between NPs and clauses	254
7.8. Conclusion	258
8 Ordering Principles, Domain Integrity, and Discontinuity	262
8.1. Introduction	262
8.2. Some morpho-syntactic properties of the languages in the sample	263
8.2.1. Word order	265
8.2.2. Relators	267
8.2.3. The coding of (properties of) referents in the predicate	268
8.3. General properties of ordering principles	270
8.4. A typology of ordering principles	270
8.4.1. Type I: placement relative to a domain	271
8.4.2. Type II: placement relative to the boundary of a domain	272
8.4.3. Type III: placement relative to a constituent in a domain	273
8.4.4. The mixed type: principles with different kinds of reference points	273
8.5. Iconicity, domain integrity, and discontinuity	274
8.5.1. Counter-examples to the Principle of Domain Integrity: discontinuity	277
8.6. Conclusion	281
9 Greenbergian Word Order Correlations and the Principle of Head Proximity	285
9.1. Introduction	285
9.2. The first hypothesis: noun-adjective adjacency	289
9.2.1. Non-adjacent adjectives in languages with a major class of adjectives	289
9.2.2. Non-adjacent adjectives in languages with a minor class of adjectives	293
9.2.3. Non-adjacent adjectives, Universal 20, and asymmetry	295
9.3. The second hypothesis: the position of modifiers relative to the noun	298
9.3.1. Word order correlations	298
9.3.2. Dryer's correlation pairs	299
9.3.3. Dryer's correlation pairs and Functional Grammar	301
9.3.4. Hawkins's principle of Early Immediate Constituents (EIC)	311
9.4. Basic word order and the position of noun modifiers in the fifty-two-language sample	318
9.4.1. Demonstratives, numerals, and adjectives	321
9.4.2. Possessive modifiers	323
9.4.3. Relative clauses	326
9.4.4. Conclusion	326
9.5. The position of embedded modifiers	327
10 The Principle of Scope	334

10.1. Introduction	334
10.2. The relative order of modifiers in non-complex NPs	335
10.2.1. Free nominal aspect markers and adjectives	337
10.2.2. Nominal aspect on adjectives	343
10.2.3. Numerals and demonstratives	344
10.2.4. Overview of ordering patterns in the simple, integral NP	347
10.3. The position of embedded modifiers	352
10.4. Conclusion	354
11 Epilogue	357
11.1. Introduction	357
11.2. Aims	357
11.3. Seinsarten	358
11.4. Nominal classification	359
11.5. Qualifying modifiers	360
11.6. Quantifying modifiers	361
11.7. Localizing modifiers	361
11.8. The underlying structure of noun phrases	362
11.9. Noun phrase internal syntax	364
11.10. Summing up	366
References	369
Index of Subjects	417
Index of Languages	433
Index of Authors	441